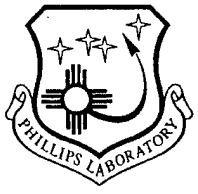
PL-TR-95-2149
Environmental Research Papers, No. 1183

PHYSICS OF THE IONOSPHERE FOR OTH OPERATION
Chapter 3, OTH Handbook

B.S. Dandekar
J. Buchau
J.A. Whalen
P. Fougere

9 November 1995



PHILLIPS LABORATORY
Directorate of Geophysics
AIR FORCE MATERIEL COMMAND
HANSCOM AIR FORCE BASE, MA 01731-3010

"This technical report has been reviewed and is approved for publication."

Maj Edward Berghorn, Chief
Ionospheric Application Branch

WILLIAM K. VICKERY, Director
Ionospheric Effects Division

This report has been reviewed by the ESC Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS).

Qualified requestors may obtain additional copies from the Defense Technical Information Center (DTIC). All others should apply to the National Technical Information Service (NTIS).

If your address has changed, or if you wish to be removed from the mailing list, or if the addressee is no longer employed by your organization, please notify PL/TSI, 29 Randolph Road, Hanscom AFB, MA 01731-3010. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document requires that it be returned.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	9 November 1995	Scientific Report

4. TITLE AND SUBTITLE
Physics of the Ionosphere for OTH operation
Chapter 3, OTH Handbook

5. FUNDING NUMBERS
PE 62101F
Proj 4643TAGH
Work Unit 01

6. AUTHOR(S)
B. S. Dandekar, J. Buchau, J. A. Whalen and P. Fougere

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Phillips Laboratory (GPIA)
29 Randolph Road
Hanscom AFB, MA 01731-3010

8. PERFORMING ORGANIZATION REPORT NUMBER
PL-TR-95-2149
ERP, no. 1183

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
This report was completed to assist the operators managing the OTH radars deployed by the US Air Force.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public release;
Distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The purpose of this report is to familiarize the Environmental Assessment radar operators with (a) the complex ionospheric phenomena that vary with location (geography), time of the day, and season and (b) with the phenomena of solar origin that affect the environment surrounding the earth. The first section describes the basic features of the ionosphere, which are an integral part of the radar system, the mirror which enables the operation of the radar. The second section discusses the effect of solar and geomagnetic activity, which control the ionospheric behavior. The last section deals with the not so uncommon disturbed ionosphere, a part of which is due to the electrodynamics of the system, and the other is to events taking place far away from the earth, on the sun. The anticipated effects of these three kinds of phenomena on the radar performance are discussed, which should help the operator to determine if performance degradation is caused by the radar system or its environment.

14. SUBJECT TERMS
OTH radars, Magnetic activity,
Auroral oval, Quiet ionosphere, Disturbed ionosphere

15. NUMBER OF PAGES
106

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

Contents

Illustrations

Acknowledgements

The authors thank Major Edward Berghorn for his valuable comments and interest in the work. The authors also thank Dr. D. C. Miller (MITRE), B. Weijers (Rome Laboratory), Carl Bowser (Martin Marietta) and Major Vincent Azzarelli for their valuable comments.

Preface

This purpose of this report is to familiarize the radar operators with the physics of the ionosphere : the ionospheric layers, the various ionospheric phenomena, and the solar and magnetic activity that control the ionospheric behavior, which is the environment that affects the performance of the radar.

We have made an effort to include most of the ionospheric physics usually encountered in the OTH operation. However the use of this information, may prove a need 1) to include additional ionospheric features and 2) provide more clarification of some of features already discussed here. Therefore we would encourage feedback and comments. Please feel free to contact:

Dr. B. S. Dandekar
Phillips Laboratory /GPIA
29 Randolph Road
Hanscom AFB, MA 01731-3010.

With the deployment of the OTH radars, a need for an OTH Handbook was realized by the community to provide a basic understanding of the OTH operation, the general layout of the radar structure, and the basic geophysics background. With the help of the OTH community, Jurgen Buchau took the lead and accepted the responsibility for producing the OTH Handbook. The planned handbook has six chapters. These are 1) Introduction, 2) OTH Radar System: System Summary, 3) Physics of the Ionosphere for the OTH Operation, 4) High Frequency (HF) Radiowave Propagation, 5) The OTH Radar Operation and 6) Glossary for OTH Radars. Chapters 2 and 4 have been published by Dr. Gary Sales as technical reports. Chapter 6 of the OTH Handbook has been published by Dr. Balkrishna S. Dandekar and Jurgen Buchau. This report forms the third chapter of the OTH handbook. The remaining chapters will be published as separate reports.

The Environmental Assessment operators (EA) have, as their major job, optimizing the radar performance at any given time in the context of highly variable ionospheric environment. To the EA operator, managing the radar operating frequency for a successful operation of the system, it could be indeed very frustrating not to understand the causes for reduced radar performance, which are specifically of ionospheric origin. However some of these problems can be reduced and in some cases the causes can be circumvented by a proper understanding of the ionospheric phenomena responsible for the degradation of the radar performance. The radar operators are more effective if they are aware of the complex ionospheric phenomena, which vary with location (geography), time of the day, and season and of phenomena of solar origin that affect the environment surrounding the earth. The first section describes the basic features of the ionosphere, which is an integral part of the radar, and which is the reflecting (/refracting) mechanism that enables the operation of the radar. The second section discusses the effects of solar and geomagnetic activity, which control the ionospheric behavior. The last section deals with the (not so uncommon) disturbed latitude dependent ionosphere, a part of which is due to the electrodynamics of the system and the other part to events taking place far away from earth, on the sun. While presenting these three aspects, we discuss the anticipated effects on the radar performance, which should help the operator to separate the performance degradation in terms of system and limitations due to environmental causes.

PHYSICS OF THE IONOSPHERE FOR OTH OPERATION

Chapter 3, OTH HANDBOOK

1. THE NORMAL IONOSPHERE

The ionized region of the upper atmosphere, nominally from 60 to 1000 km above the surface of the earth, plays a critical role in the Over-the-Horizon radar operations. These radars operate in the frequency range 3 to 30 MHz, known as the HF band. The radiated radar signal is refracted (reflected) by the substantial number of free electrons in the ionosphere, back down to the ground and to a target, flying above the ground several thousand kilometers from the radar (see report by G. Sales PL-TR-92-2123 for a detailed description of the interaction of the radar signal with the ionized region of the atmosphere). The signal reflected from the ground/ target is used by radar for target information (correlation, tracking).

The ionosphere is considered an integral component of the OTH radar system. Without the ionosphere, there is no "over-the-horizon" radar. On the other hand, the complexity of the operation of these radars, as compared to the usual line-of-sight radar systems, results from the complex and unpredictable behavior of the ionosphere. Optimum OTH performance depends upon an accurate and timely representation of ionospheric behavior using the radar system's software and a real-time ionospheric and geophysical data base to generate a specification of the relevant ionosphere; that is, a description of the distribution of the electrons as a function of altitude, longitude, and latitude over the radar's coverage region. The ability of the OTH radar to precisely position and identify targets depends upon the accuracy of the description of the ionosphere in both space and time.

Received for Publication 9 November 1995.

Ultimately, the sun is responsible for the ionosphere's most important characteristics. Figure 1 shows a simple schematic flow chart that describes the sun-earth interconnection. The solar X rays and Extreme Ultra-Violet (EUV) radiation creates the ionosphere. The activity of the sun (see later in this section), and the interaction of the solar wind/ interplanetary magnetic field (discussed later) with the earth's magnetic field which results in geomagnetic activity, affects the ionosphere and makes the prediction of the changes in the ionosphere very difficult. The OTH radar operator, in response to the ionospheric changes, must adjust the radar's operating parameters (primarily the operation frequency) to operate the radar effectively. In the absence of all solar activity and of perturbations in the solar-interplanetary-earth magnetic fields, it would be reasonable to expect the ionosphere to behave in a very predictable manner. Such a "quiet" ionosphere is, in fact, what is seen when climatological ionospheric data are examined, such as the mean electron density profiles (electron density vs altitude) shown in Figure 2 for day and night and at two times, approximately 5 years apart, in the 10.5 year cycle of solar variations. In fact, the solar-interplanetary-earth environment is highly variable over relatively short time scales (hours and days) is never truly "average" and never remains quiet for very long. This variability makes the short term prediction of ionospheric behavior difficult if not impossible, despite the great strides made in the last 30 years in understanding the solar-terrestrial relationship.

1.1 Vertical Structure

The complexity of the basic ionospheric structure is best illustrated in Figure 2 showing the altitude variation of the electron density with the apparent formation of ionospheric layers, and in Figure 3 which illustrates the global variations of one of the most important ionospheric parameters (f_oF_2), directly related to the maximum electron density often found in the 250 to 450 km region above the earth's surface.

The layered structure in the electron density is identified in terms of the D, E and F layers or regions, each representing change in the effects of the solar ionizing radiation on the changing neutral atmospheric constituents as the altitude increases. The ionosphere always remains charge neutral, that is positive and negative charge are always created in pairs and though they may

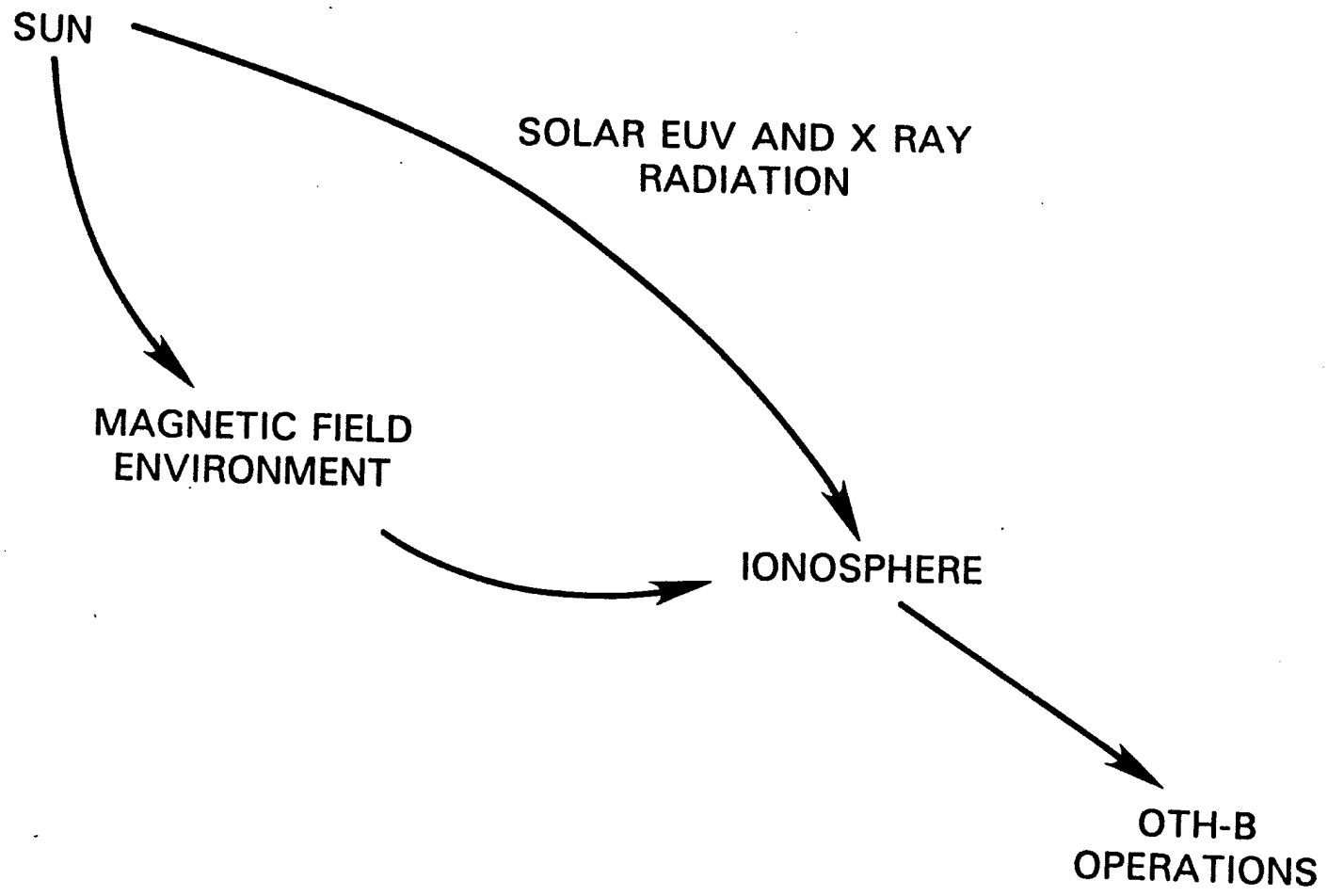


Figure 1. The sun earth connection affecting the OTH radar operation.

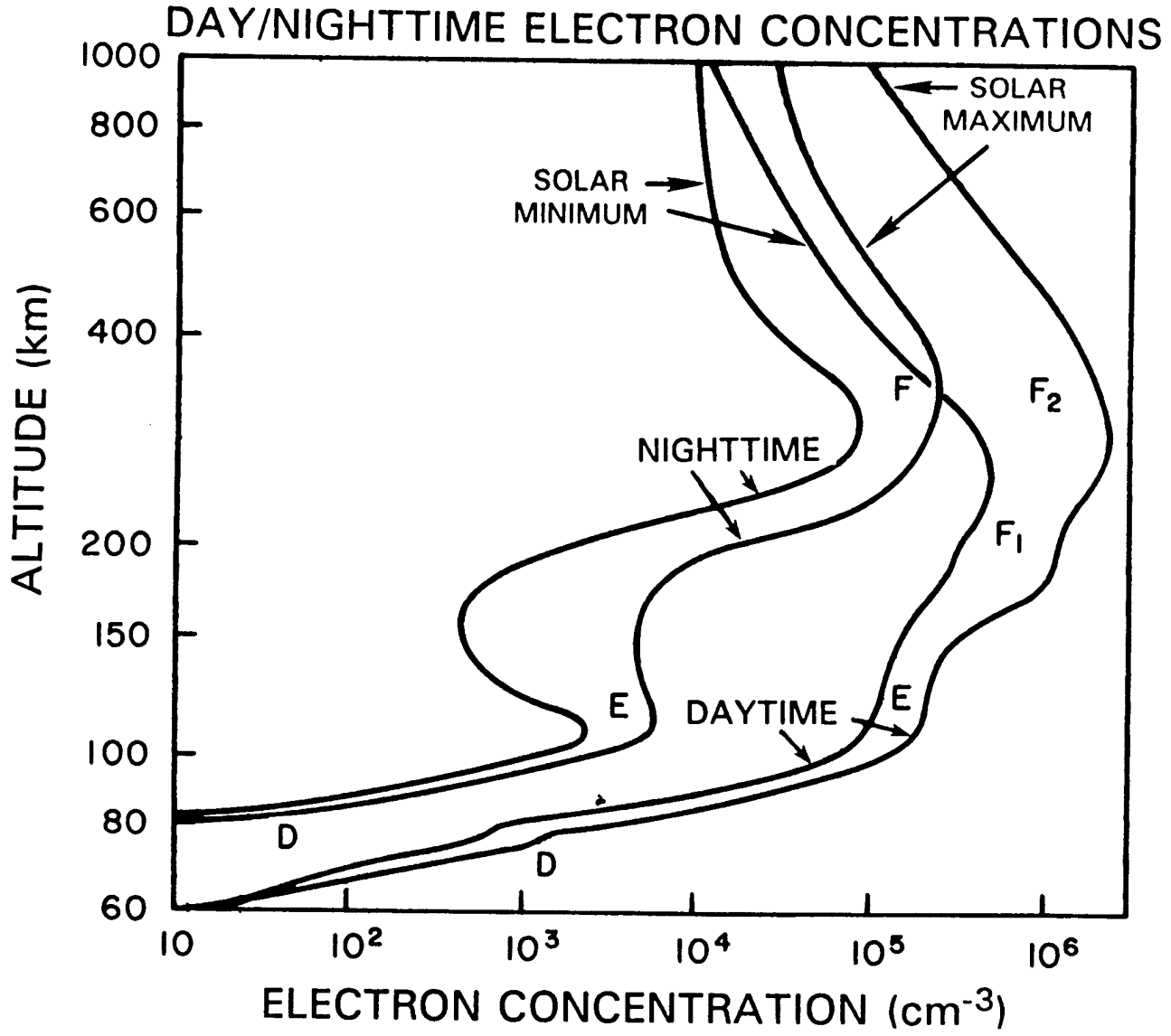


Figure 2. Layer structures and electron density profiles for day and night for solar maximum and solar minimum periods, (Jursa[1], p.9-1).

distribute differently at different heights, there is always one positive charge for each negative charge. It is the F-layer (the highest and in terms of electron density strongest layer) of the ionosphere that plays the dominant role in the propagation of the radar signals to over-the-horizon distances. It is also essential to understand the behavior of the lower ionospheric layers as they also affect HF radio propagation (for example in summer daytime E layer provides a 2 hop propagation mode) and therefore the management and performance of the OTH radar. The variability of the various layers from day to night and with solar cycle is evident and will be discussed later.

The global variations in the maximum electron density are as complex as the vertical structure. The particular map shown in Figure 3, is for 22 UT in January, and for a moderately high sunspot number of 110, and presents the changes from the sunlit ionosphere to the dark hemisphere. Superimposed on the diurnal variation, is a strong latitudinal dependence, influenced by the changing solar zenith angle (the angle the incoming solar radiation makes with the vertical direction at the observation point), the inclination of the earth's magnetic field to the vertical direction (dip) and by solar wind particles (electrons, protons and alpha particles) which are able to enter the earth's atmosphere in the auroral zones and polar caps, where they also ionize the ambient medium.

The vertical structure of the ionosphere varies continuously during the day, with the seasons of the year, with latitude and with the phase of the solar cycle. It is sensitive to enhanced periods of short, energetic short wavelength solar radiation (X-rays and EUV) which occurs during solar flares, and to the more sustained effects of geomagnetic storms. (The discussion of solar flares and geomagnetic activity follows later in this section). In spite of all this, the essential features of the ionosphere are usually identifiable, except during periods of intense solar/ geomagnetic disturbances. The different horizontal layers in the ionosphere are shown in Figure 2. The D, E, F_1, and F_2 layers are shown; increasing in altitude and in electron concentration. Figure 2 also shows typical midlatitude daytime and nighttime vertical electron density profiles as they change over the course of the 10.5 year solar activity cycle (solar activity is measured by the numbers of sunspots on the solar disc, with few sunspots representing quiet, and 150 sunspots active conditions). Above the maximum electron density of the F_2 region, the electron density steadily decreases out to several thousand kilometers above the surface of the

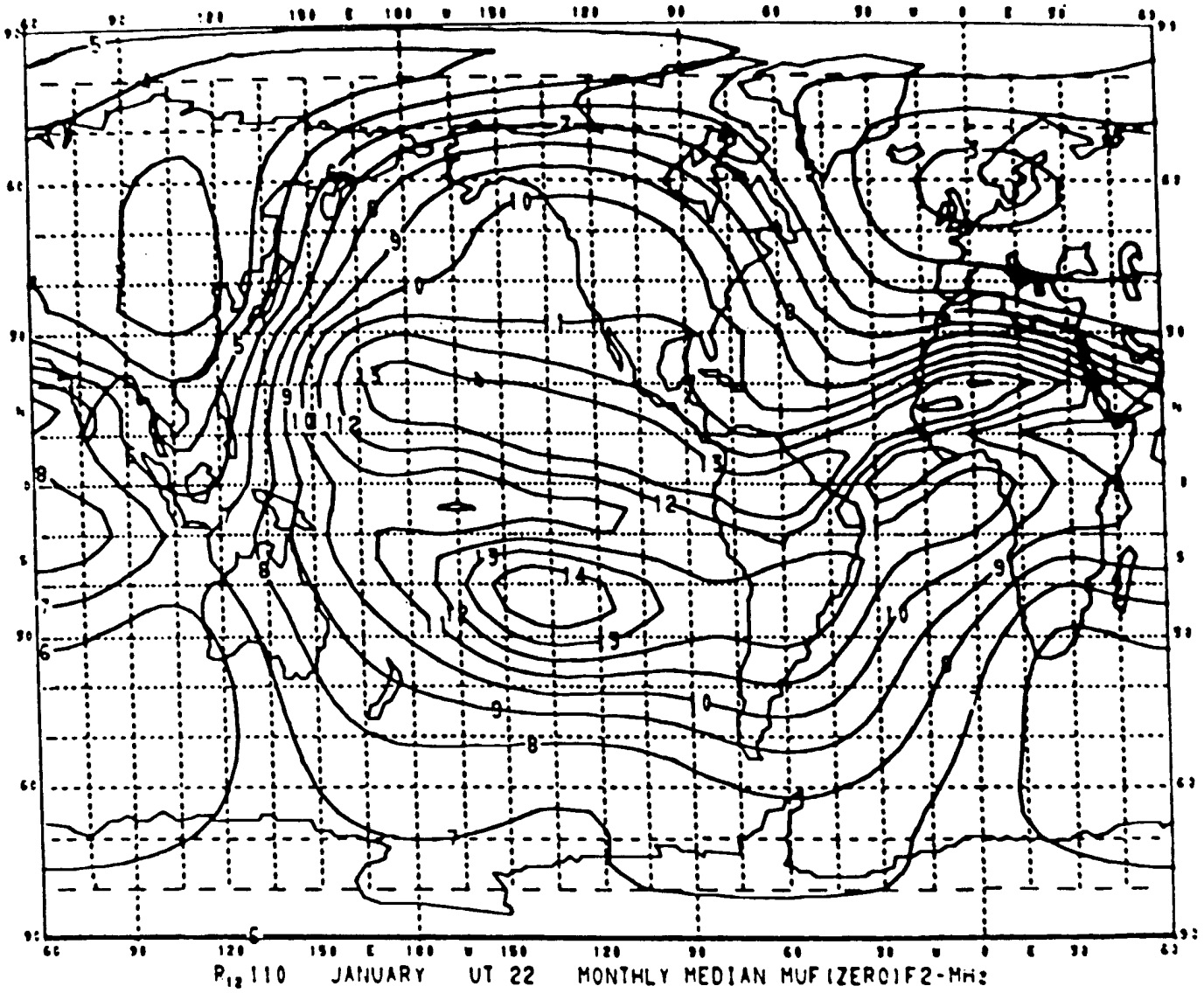


Figure 3. Global f_oF_2 contours for the month of January at 22 UT and SSN=110, (Roberts and Rosich[2], p.25).

earth. The region above about 1000 km is called the protonosphere (since it is populated predominantly by ionized H atoms, that is, electrons and protons) and as will be discussed later, supports the maintenance of the night time F layer at midlatitudes. During the night the electron density decreases, since there is no solar production, with electron losses occurring due to recombination of the electrons with positive ions. In fact, the F_1, E and D layers disappear for all practical purposes soon after sunset, leaving only the F layer.

The formation of distinct ionospheric regions, the D, E, F_1, and F_2 layers, results from three primary factors:

a. The ionizing solar radiation deposits its energy at various heights depending on the absorption characteristics of the atmosphere.

b. The rate of recombination (of electrons and ions) depends on the atmospheric density, which changes with height.

c. The chemical composition of the atmosphere changes with height.

The details of the formation of these regions is discussed in Section 1.3, but suffice it to say at this point, that the electron density at any altitude is a balance between the production of electrons by ionization and the loss of electrons by chemical recombination of electrons and ions, and by diffusion upward and downward.

In the D and E regions from 60 to 150 km altitude the production of electrons by the ionization results in O_2^+ and NO^+ as the dominant molecular ions. The loss process is called dissociative recombination where the electrons combine with the O_2^+ and the neutral product is atomic oxygen ($O_2^+ + e \rightarrow O + O$). The electron density in these lower layers follows closely the intensity variations of the solar radiation, therefore also following closely the change in solar zenith angle. As a consequence these three layers have their maximum densities at local noon. We will talk about this in more detail later.

In the ionospheric F_2 region, O^+ is the dominant ion and loss is through a charge exchange process with N_2 and O_2 ($O^+ + O_2 \rightarrow O_2^+ + O$) followed, as above, by the dissociative recombination process. The loss rate for O^+ ions is very low at the F region altitudes since the number of the heavier molecules, N_2 and O_2, decreases with altitude. This very low loss rate means, that at night, the electrons can persist for many hours after the sun has set, thereby, providing the propagation medium to support OTH operations, even at night when the sun does not shine. The low loss rate also means that transport processes including diffusion and wind

motions can affect the electron density distribution and complicate the simpler behavior of the lower ionospheric layers which are under direct solar control.

It is common practice, in the radio and OTH communities, to speak of the ionospheric layer densities in terms of a plasma frequency, rather than electron density. The plasma frequency f_p of a given electron density is equal to the radio frequency, which a stratified slab at that density reflects at vertical (perpendicular) incidence. The electron density N_e is related to the plasma frequency by the relation:

$$N_e = 1.24 \times 10^4 f_p^2$$

where N_e is the peak electron density (in cm^{-3}) and f_p is the plasma frequency (in MHz).

The maximum plasma frequency of a layer, called the critical frequency of the layer, is equal to the highest radio frequency transmitted vertically that is reflected from the layer. The critical frequencies are denoted with f_o, followed by the respective layer identifier. The critical frequencies of the E, F_1 and F_2 layers are therefore f_oE, f_oF_1, and f_oF_2, respectively. At incident angles other than vertical, a given layer can reflect frequencies higher than the critical frequency. (There are useful empirical relations between the critical frequency and height of a layer, angle of take-off, and the radio frequencies propagated via that layer to a given distance). This is discussed by G. Sales in his report (PL-TR-92-2123) on radiowave propagation in the ionosphere.

As a useful reference, the following table relates typical electron densities and their respective plasma frequencies.

Table 1.

N_e (cm^{-3})	10^4	10^5	10^6
f_p (MHz)	1	3	10
Typical for	E Layer (sunrise)	E Layer (noon)	F_2 Layer (daytime)

Electron density is given in this report in terms of electrons per cm^3, shown in scientific notation as cm^{-3}.

1.1.1 D-LAYER

This lowest of the identifiable ionospheric regions is located in the height range from 50 to 90 kilometers. It is produced by the most penetrating of the solar ionizing radiations. Electron densities are small, ranging from 10^2 to 10^4 cm^{-3} in the day to nearly zero at night, except at the high latitudes, where particle precipitation can maintain ionization, even at night. The diurnal variation dominates both the seasonal and solar activity cycle dependence in the D-region. It is this region of the ionosphere that affects the attenuation of the radio waves passing through this height range, which the radar signal must traverse twice on the outgoing leg and twice again on the return leg of the propagation path (thus over all, the layer is traversed a total of four times). The D-region attenuate HF radiowaves because of the high rate of collisions between the electrons and the very dense neutral atmosphere at these low altitudes. The radio wave forces the electrons to oscillate with the wave frequency. With each collision the radio wave loses a little bit of its energy. Therefore, the more electrons existing at D-region altitudes, the more energy is extracted from the radio wave. The D layer absorption is known as non-deviative absorption.

1.1.2 E-LAYER

The E-layer is located between 90 and 140 km altitude and has a maximum at 100 km with a number density near 10^5 cm^{-3} corresponding approximately to a critical frequency of f_oE = 3 MHz. The E layer is strongest during sunlit hours, local summer, and at solar activity maximum (see f_oE plot in Figure 4). High nighttime E-layer ionization densities occur at higher latitudes, where particle precipitation penetrates and maintains the electron density (auroral E). In the absence of the particle precipitation the E layer shows a good solar control, with f_oE given by the equation

$$f_oE(MHz)= 0.9*[(180+1.44*R)*\cos\chi]^{0.25}$$

where R is the monthly median running average sunspot number and χ is the solar zenith angle. Using a highest sunspot as 150, the highest magnitude of f_oE rarely exceeds 4 MHz. This layer dissipates soon after sunset. The daytime E-layer can support the propagation of radar signals and often, particularly during the summer period, provides the best propagation modes for over-the-horizon target detection, though at shorter ranges (usually from 1000 km to 2000 km) than detection provided by the higher layers.

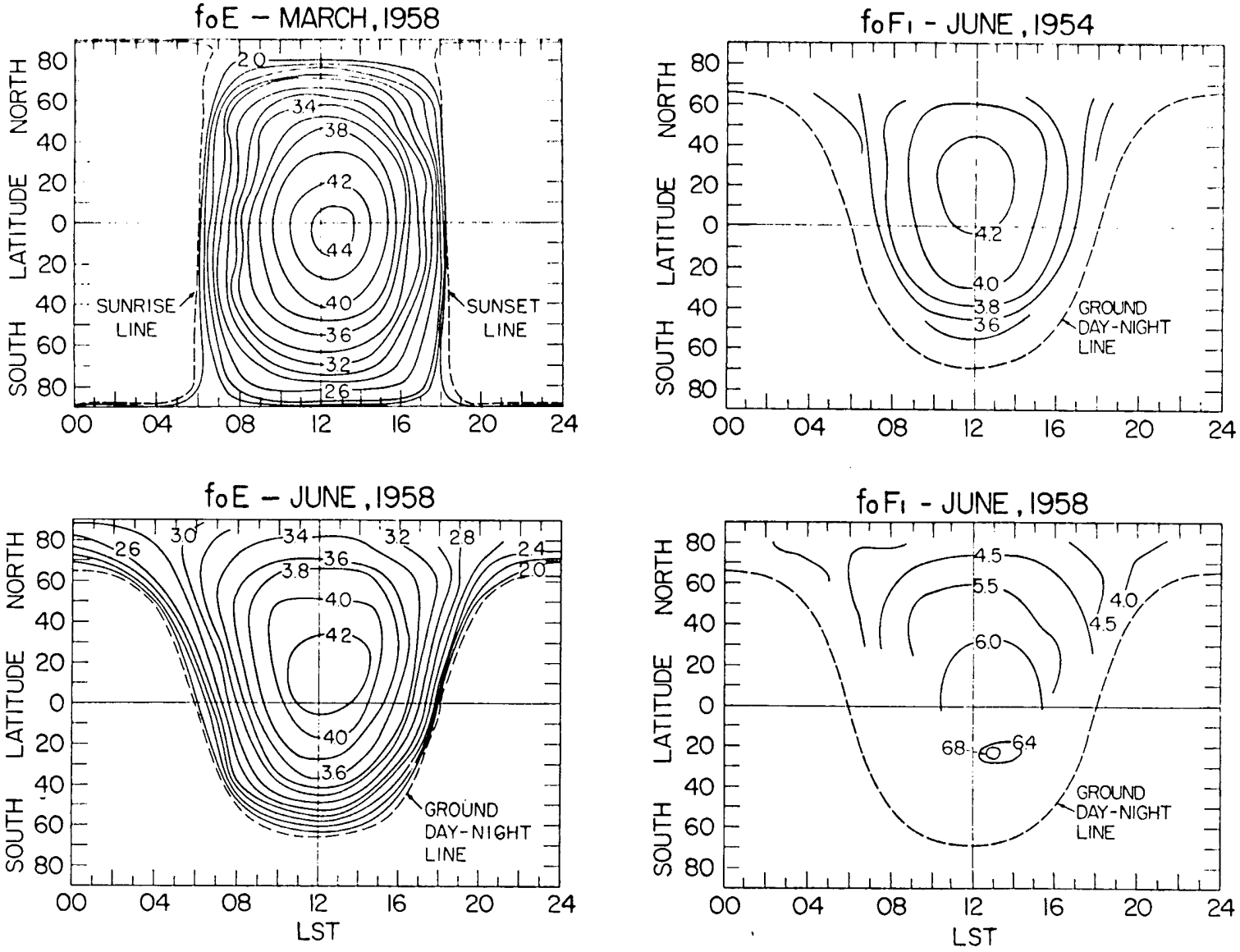


Figure 4. Spatial, temporal, seasonal and solar activity dependence of f_oE and F_1 layers, (Davies[3], p.139 and 141). Contour lines are frequency in MHz.

1.1.3 F_1 LAYER

The distinction between the F_1 layer and the F_2 layer exists only in the daytime, and very frequently, the difference is ignored and the entire region is referred to as the F layer. Located around 140 to 200 km height, the F_1 layer is a daytime ledge at the bottom of the F_2 layer, with maximum electron density around 3 x 10^5 (cm^{-3}) or f_oF_1 = 5 MHz (see f_oF_1 plot in Figure 4). Similar to the E layer, the F_1 layer shows solar control with f_oF_1 given by the equation

$$f_oF_1(MHz) = (4.3+0.01*R)*\cos^{0.2}\chi$$

where R is the monthly median running average sunspot number and χ is the solar zenith angle. Like the lower E layer, it is most pronounced in summer and during sunspot maximum. The F_1-layer can often support radar signals, though it is difficult to separate these modes in backscatter sounding from the F_2 modes.

1.1.4 F_2 LAYER

This layer has the largest electron densities in the ionosphere. Though this layer is solar produced like the E and F_1 layers, the transport processes strongly dominate the behavior of this layer. Its height varies greatly: 225 km on winter days to 400 km on winter nights at midlatitudes and 25 to 100 km higher in summer and during solar cycle maximum. At low latitudes the height may vary from 300 to 600 km depending on local time and phase of the solar cycle. Day/ night electron densities are of the order of 10^6 and 10^5 cm^{-3}, or f_oF_2 = 10 and 3 MHz, respectively, showing a large degree of variability. The highest densities are during sunlit hours in winter at solar maximum. The fact that these highest densities (critical frequencies) occur during the winter season indicates, as discussed earlier, that solar control is not the only factor that affects the distribution of electrons. The reason for this "winter anomaly" will be discussed later. Transport phenomena are also very important in the F-region and make the prediction of its behavior much more difficult. There also exist other seasonal and latitudinal anomalies in the F-layer electron density distribution. As stated earlier, the F-region is the most important ionospheric region, from the point of view of OTH radar propagation management, since it provides the reflection region for long range (>2000 km) propagation (see report by G. Sales PL-TR-92-2123).

1.1.5 SPORADIC E

Localized patches of increased ionization frequently occur at altitudes of the normal E-region regime. Because of the unpredictable nature of the specific time and location of these patches, the phenomenon is called sporadic E (E_s). On a statistical basis, midlatitude E_s is a summer daytime phenomenon. Sporadic-E is characterized by vertical thickness of a few kilometers, horizontal expanse of several hundred kilometers, and durations of several hours. Equatorial E_s, closely related to equatorial E region currents, occurs about 80 percent of the time during daylight hours, with no seasonal preference. Sporadic E in the midlatitudes (temperate E_s) is probably related to horizontal wind shears acting on the local plasma in the presence of the earth's magnetic field. Temperate E_s is primarily a daytime, summer phenomenon, occurring 30 to 70 percent of the time, with the lower probabilities occurring at higher latitudes. The critical frequencies of these E_s layers are highly variable, with a median value of around 5 MHz, and not infrequently, reach 7 MHz. These intense E_s occurrences blanket the F region, making F-mode propagation impossible. When this occurs, the use of E-mode propagation markedly decreases the range at which the OTH radar can operate, though performance can be quite good. Maximum useable range for the 1 hop E_s mode is 2000 km (see report by G. Sales PL-TR-92-2123), and at times E_s may support the two-hop mode.

1.2 Latitudinal Structure

The structure of the electron density distribution in the quiet ionosphere varies greatly with latitude. This north/south variation is better organized by a geomagnetic coordinate system than by the geographic coordinate system. The geomagnetic coordinate system is a latitude/ longitude coordinate system centered on the axis of the earth's magnetic field, which is displaced from the earth's rotational axis that defines the geographic coordinate system. The difference between the two systems is best seen by overlaying the geomagnetic latitude system on an ordinary geographic map (Figure 5). The latitudinal structure of the ionosphere is controlled by electrodynamic processes that interact with the earth's magnetic field and therefore show a strong geomagnetic latitude dependence. The strong geomagnetic control of the ionosphere can also be seen in the global f_oF_2 map (Figure 3), which shows f_oF_2 in a geographic coordinate system. The lowest f_oF_2 values between the two equatorial irregularity regions swing from the geographic equator at 180° W down to near 10° S near 90° W (near South America), and up to

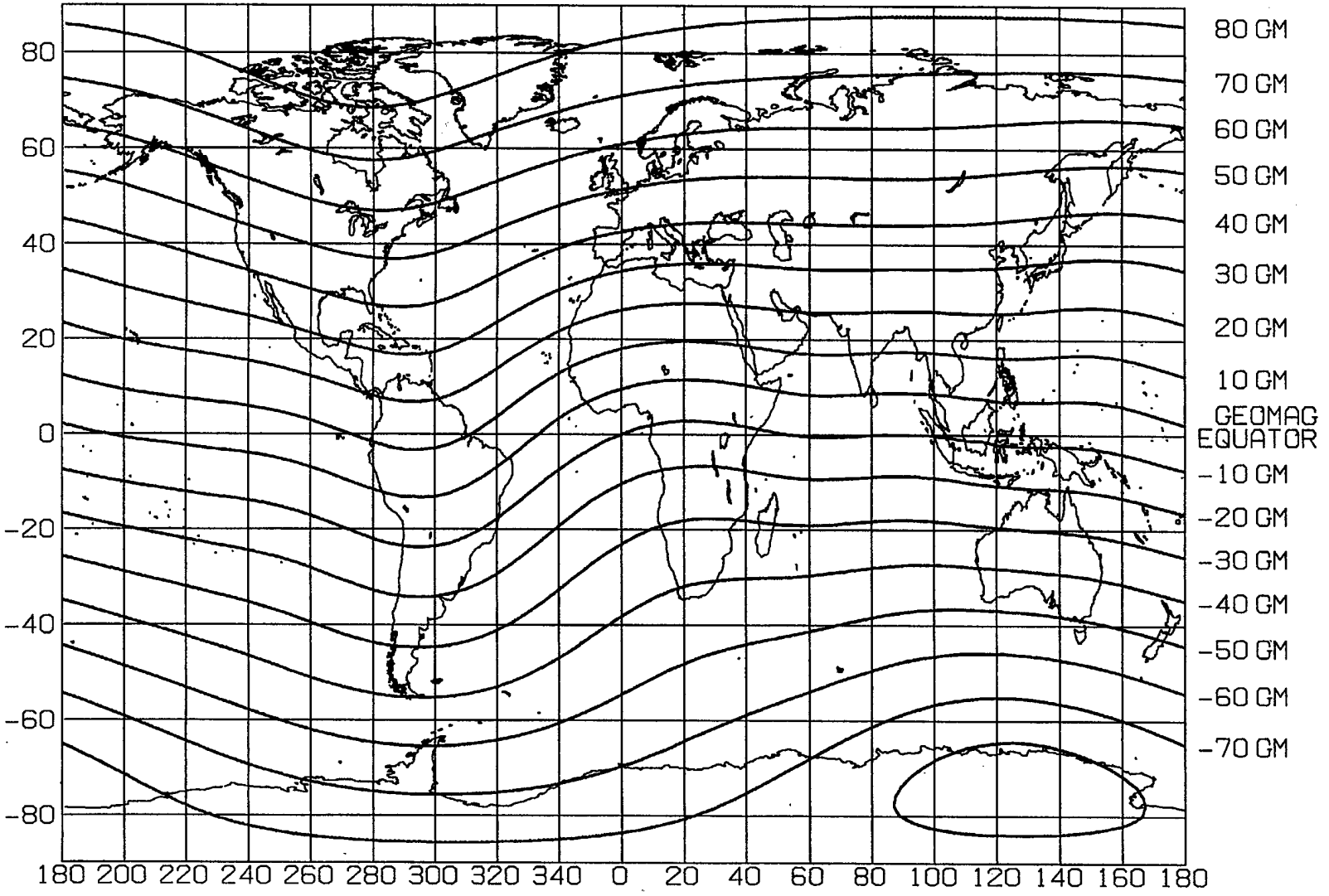


Figure 5. Geographic and Corrected Geomagnetic coordinates for earth.

10° N over Africa, near 0° longitude, closely following the path of the magnetic equator. To a lesser extent, atmospheric winds (transport mechanisms) also affect the configuration of the ionosphere and introduce a geographic component to the distribution of electrons.

The ionospheric behavior is better ordered in geomagnetic than in geographic coordinates. The global ionosphere can be divided into three latitudinal regimes; low latitude (0° - ±25° geomagnetic), middle latitude (25° - 55° in two bands north and south of the equator), and high latitude (55° - 90°), also in the northern and southern hemispheres (see figure 5). Of these regions, the midlatitude ionosphere is the most well behaved, which is a fortunate circumstance for the OTH operation.

1.2.1 LOW LATITUDE REGIME (EQUATORIAL IONOSPHERE)

At the equator, the magnetic field is nearly horizontal to the surface of the earth. Because of this special geometry, there is an intense current sheet, known as the equatorial electrojet, which flows along the magnetic equator at about 110 km altitude (E region) in a strip a few degrees wide in latitude. The current flows toward the east in the day where the ionospheric density is high in the E region. There is a westward flow at night, but this is nearly undetectable because of the small electron concentrations present at nighttime.

The electric fields associated with these ionospheric currents drive a plasma convection in the F region at low latitudes that is upward and westward in the daytime. This upward motion takes freshly ionized plasma to higher altitudes where it recombines very slowly. This plasma which has moved up, now flows under the influence of gravity down the magnetic field lines to higher latitudes (± 20°) both north and south and results in regions of increased electron concentration in two bands parallel to the geomagnetic equator. This phenomenon gives rise to the so-called "equatorial or Appleton anomaly" (lower electron density at the equator and higher electron density at ± 20° latitudes). As a result of these and other physical processes, daytime critical frequencies for all equatorial ionospheric layers are generally higher than at midlatitudes. There is little particle precipitation at low latitudes so that in this region ionization is produced primarily by solar EUV radiation.

Enhanced upward transport just after sunset causes the F-layer to reach altitudes as high as 600 km near the geomagnetic equator with a strong propensity to become unstable and generate ionospheric irregularities (see discussion later in this Section and in report by G Sales PL-TR-92-2123) which result in considerable latitudinal variation in the electron density distribution in the premidnight hours, severely disturbing radio and radar systems operating along paths that intersect this region. Other low latitude disturbances that are not generally related to geomagnetic or solar disturbances are of manmade nature (such as industrial machinery).

1.2.2 MID LATITUDE REGIME

The mid latitude ionosphere (25° to 50° latitude) best represents what might be called the "classic" ionosphere. The midlatitude region is free of the direct influence of the geometry of the horizontal magnetic field that is present in the equatorial region. Also, this area is generally free of the direct influence of energetic particle precipitation associated with the higher latitudes.

The midlatitude F-region is fairly well understood. In the F_2 region, the electron concentration is highest during the daytime, while during the nighttime the electron concentration decreases by a factor of 10. The nighttime F-region is maintained by atmospheric wind patterns and by a downward flow of charged particle plasma from the protonosphere (the high altitude region above 1000 km where H^+ ions predominate). The altitude of the F_2-layer peak is generally lower in daytime than during the night.

Midlatitude F_2-region electron densities, when expressed in terms of the critical frequency, f_oF_2, respond linearly to solar activity as measured by sunspot number (defined on page 29); the F_1 and E-regions respond similarly, but with a smaller variation (Figure 6). This is reasonable, since the increased flux of EUV rays during the high solar activity period is most effective in producing ionization at the higher levels, in the F-region, where the neutral atmospheric density is also highest at solar maximum.

There are several anomalous phenomena in the midlatitude electron density; the most well known is called the winter anomaly. One would expect the greatest electron densities in the F_2-region to occur in summer when the sun is more directly overhead. Instead, the seasonal maximum occurs in winter (Figure 7). This effect is related to changes in the neutral

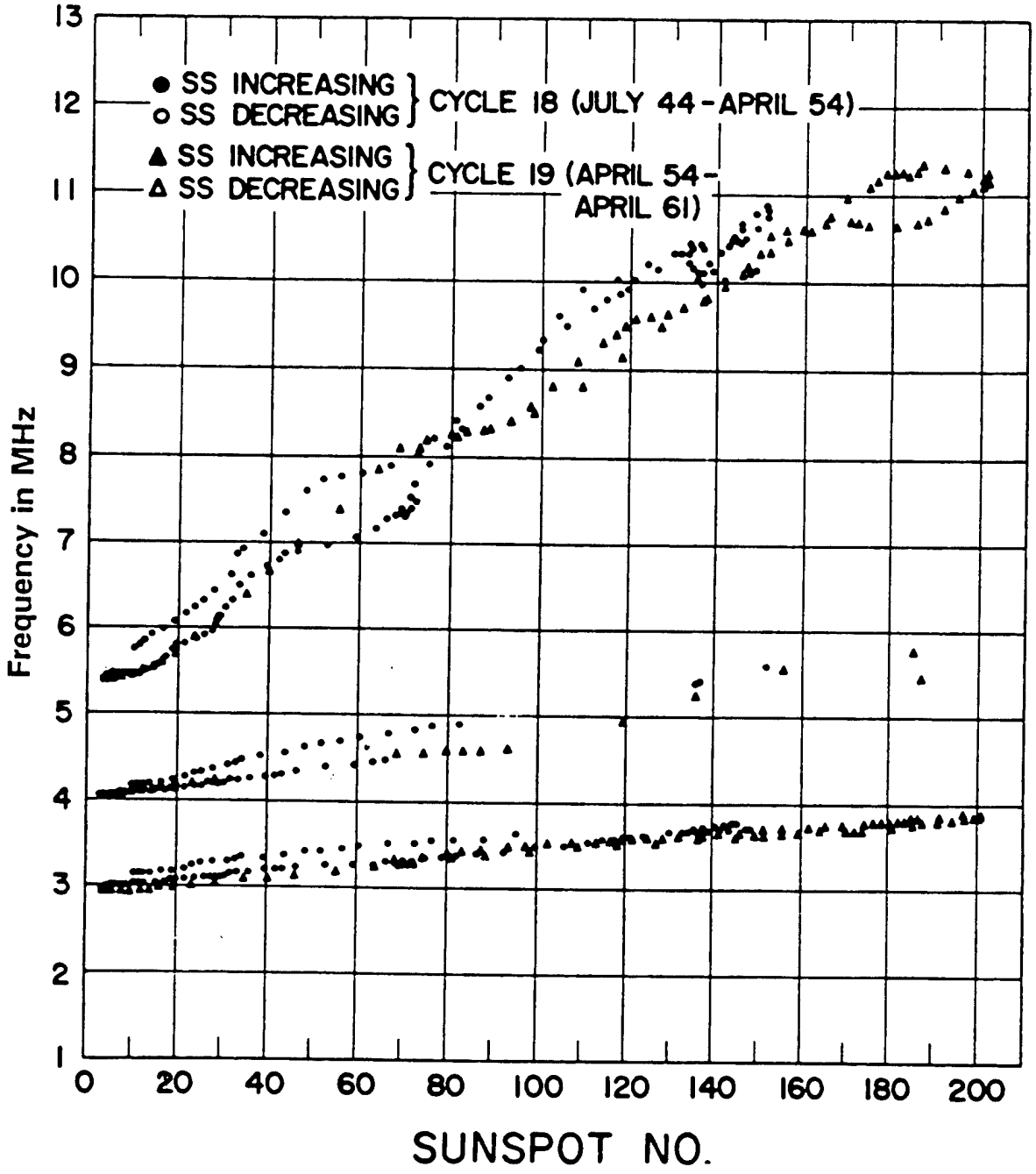


Figure 6. Critical frequency dependence of E, F_1, and F_2 layers for solar cycles 18 and 19, (Davies[3], p.144).

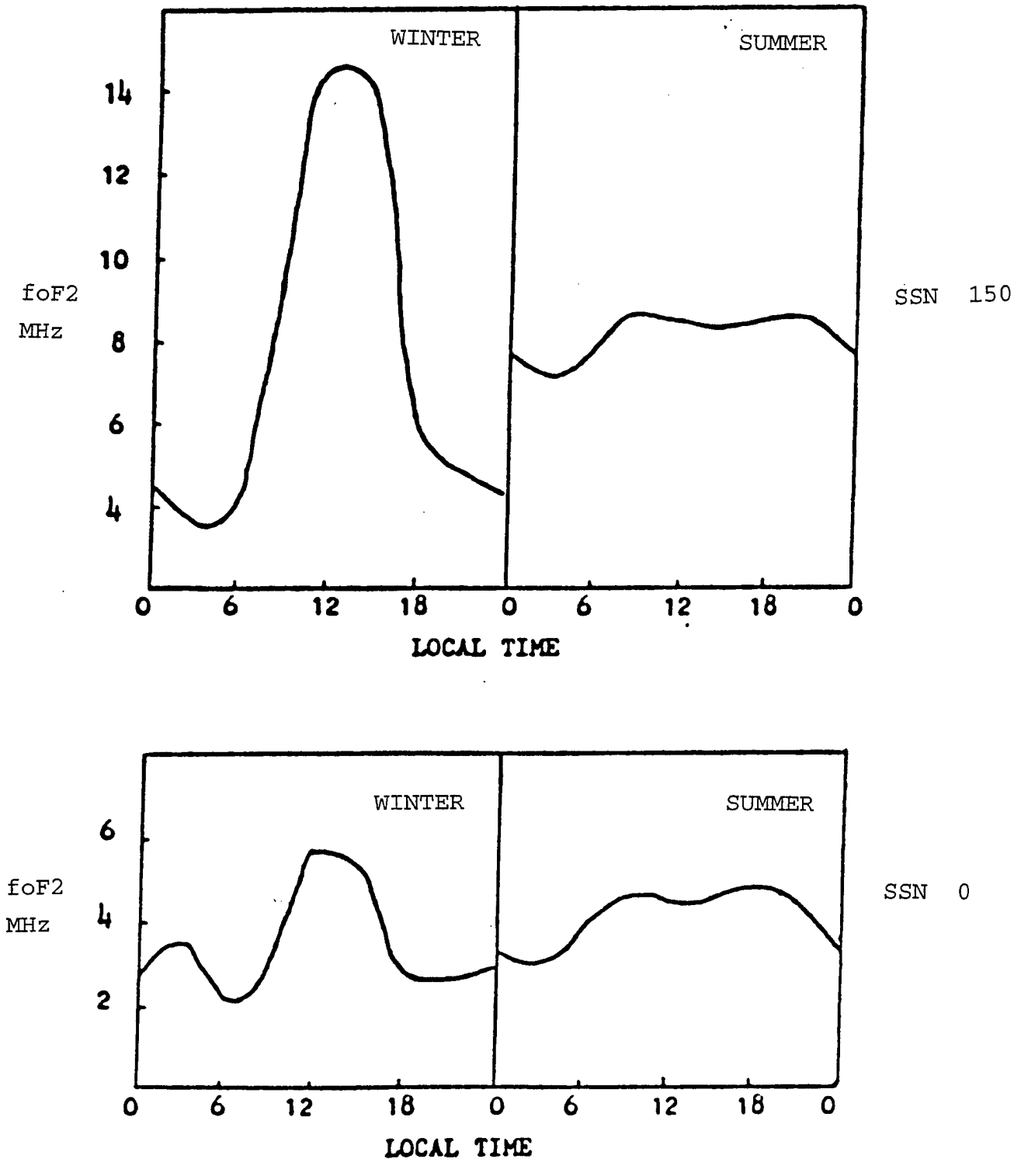


Figure 7. Diurnal, seasonal and solar activity dependence of f_0F_2 layer.

atmospheric densities which in turn affect the electron-ion loss rate; the winter anomaly is most noticeable during the solar maximum period.

The middle latitude ionosphere has somewhat lower densities than the equatorial ionosphere, as far as the F-region is concerned. The D and E-regions, on the other hand, remain under strict solar control from the equator to ±55° latitude. The middle latitudes are characterized by anomalies, such as the winter anomaly, as well as strong variations with solar activity, time of day, and season. However, these reasonably systematic variations are very well understood and are rather predictable. These can be called systematic irregularities of the undisturbed ionosphere.

1.2.3 HIGH LATITUDE REGIME

The high latitude region covers global regions typically poleward of 55° geomagnetic latitude. This region is generally under the influence of energetic auroral particle (electrons and protons) precipitation, and of strong electric fields generated by the interaction of the solar wind with the outer reaches of the earth's magnetic field. It is significantly different from the middle and low latitudes. Here, the precipitation of energetic particles is an important source of ionization, particularly at the height of the lower layers such as the D and E-regions. Transport of ionization produced in the noontime ionosphere below the auroral oval into the high latitude region by the high latitude electric fields is also an important factor in the maintenance and redistribution of the high latitude ionosphere.

The high latitude region is divided into three areas, the polar cap, the auroral oval, and the subauroral or midlatitude F-region trough. The auroral oval, a band of auroral forms encircling the magnetic pole at all times, is an active region of particle precipitation and electric currents. The most obvious characteristic of the oval is the visible light emissions, the auroras, which are observed even during quiet (undisturbed) periods. Most of the particle precipitation that creates the discrete, visible aurora occurs in the nighttime sector. In the day sector the aurora is weaker and has less latitudinal extent. Superimposed on the rather erratic precipitation leading to the spectacular displays is a steady spatially uniform band of precipitation, the continuous aurora. This continuous aurora exists even at quiet conditions, although with less intensity and at higher geomagnetic latitudes than the disturbed auroral oval. Its optical signature is a uniform, barely visible glow; its ionospheric signature is the auroral E layer, a band of E region ionization along

the auroral oval. Due to currents flowing in this medium, the auroral E layer is heavily structured and therefore a potential constant source of auroral clutter to OTH radars even under quiet conditions.

The Space Forecast Center (SFC) routinely provides the OTH operations centers with a Q index, which describes the geomagnetic latitudinal extent of the auroral oval. The Q index and its relation to the auroral boundaries will be discussed later in this section. The DMSP (Defense Meteorological Satellite Program, see Pike[4]) satellites help determine this index by providing photographs of the aurora, as well as J/4 sensor data, which show the latitudes at which precipitating electrons are being detected.

Poleward of the auroral oval is the polar cap. In winter a large part of this region is in continuous darkness and the ionosphere is maintained both by transport of solar produced plasma by solar wind driven convection (geomagnetically active conditions), and to a lesser extent by soft (less energetic than in the oval) particle precipitation (geomagnetically quiet conditions). Both convected and particle produced plasma in the polar cap are highly structured. This structured medium is transported into the night sector of the auroral oval and into the suboval night-time trough and is a contributing cause to the strong spread clutter observed by the OTH radar in the northerly direction (the other cause for the clutter is irregularities resulting from the particle precipitation into the auroral oval, and strong local plasma shear flows along auroral arcs).

Equatorward of the night-time oval is the subauroral or midlatitude F-region trough. The trough is a region of markedly reduced electron concentrations, especially in the winter nighttime. The trough is believed to be caused by increased recombination as a result of strong convection, and by transport of low density night-time (polar cap) plasma along the flanks of the auroral oval towards and (under moderately active conditions) beyond the day/night terminator (daytime trough). The trough, due to its low electron densities (f_oF_2 <= 2 MHz) and its structured ionization, is a most difficult region for OTH operation. Figure 8 shows the locations of the auroral E layer and F layer trough along a latitudinal section in the earth's noon midnight meridian plane.

Predicting radar signal propagation characteristics in the high latitude region is difficult even during quiet solar-terrestrial periods. During disturbed conditions, as we will see in the next section, it is very nearly impossible!

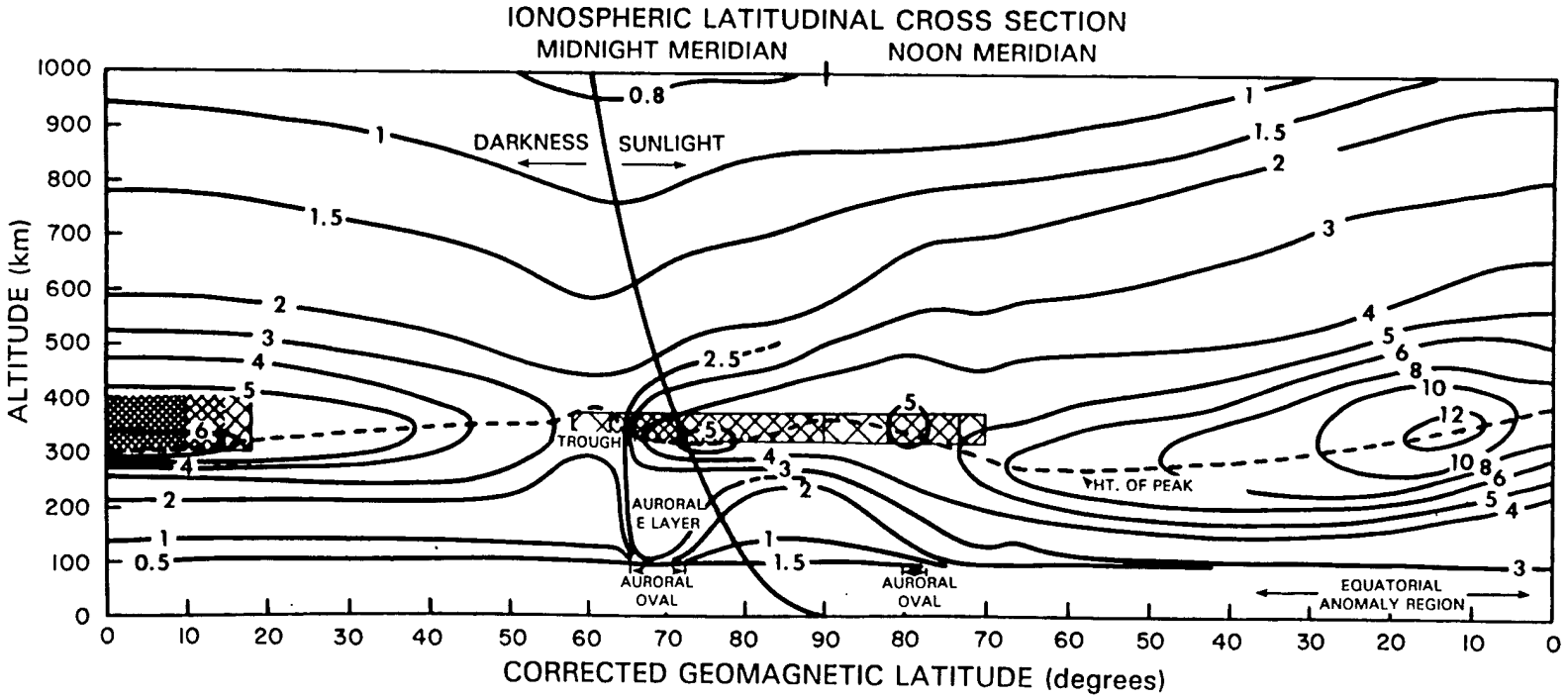


Figure 8. Schematic cross section of the ionosphere in the noon- midnight meridian plane, across the magnetic pole for solar minimum, equinox, 1200 UT. The contours show plasma density, which is a measure of electron density.

1.3 Formation/ Maintenance of the Ionosphere

The two major ionizing sources that are responsible for the structure of the earth's ionosphere are: (1) solar radiation, and (2) particle precipitation. Solar X-rays (1 to 10 A°) and extreme ultraviolet radiation (EUV between 100 to 1000 A°) are of major importance at low and middle latitudes, as these photons are sufficiently energetic to photo-ionize some of the constituents of the high atmosphere. These solar radiations are part of the normal output of the sun, though they represent only a small fraction of the total flux (the predominant part of the sun's output is in the visible portion of the spectrum; these photons cannot ionize the neutral atmospheric constituents). At higher latitudes, on the other hand, there is, in addition to the solar radiation, a steady stream of solar wind charged particles (mostly electrons) which spiral into the upper atmosphere along the earth's nearly vertical magnetic field lines, causing ionization by collision. Each of these two ionization sources are examined separately.

1.3.1 SOLAR RADIATION

Figure 9 shows solar energy output as a function of wavelength. The visible light lies between 3000 A° and 6000 A° The portion of the spectrum responsible for photo-ionizing the earth`s upper atmosphere (wavelengths below 1800 A°) represents a very small fraction of the total solar output (only about 1 part in 10^7 of the energy in the total spectrum). The amount of ionization produced in the upper atmosphere is very sensitive to changes in this portion of the spectrum, even though the change in the total solar output is extremely small. Significant changes in the ionizing portion of the spectrum occur during solar flares, when the X-ray/EUV flux may increase for several minutes by a factor of 10 to 1000. The ionosphere responds to changes in this portion of the solar radiation spectrum in a variety of ways, whether these changes are due to special flare activity or simply diurnal changes in solar zenith angle.

High energy, ionizing radiation penetrates into the earth's atmosphere to a depth that depends on the solar zenith angle. When the zenith angle is small (near noon and/or at low latitudes), the incoming radiation encounters fewer atmospheric constituents per unit depth, as compared to radiation arriving at a more shallow angle, and gives up its ionizing energy more slowly than for shallow angles, thus penetrating to lower altitudes. As the solar zenith angle becomes larger (the sun becomes more shallow), the radiation encounters more atmospheric

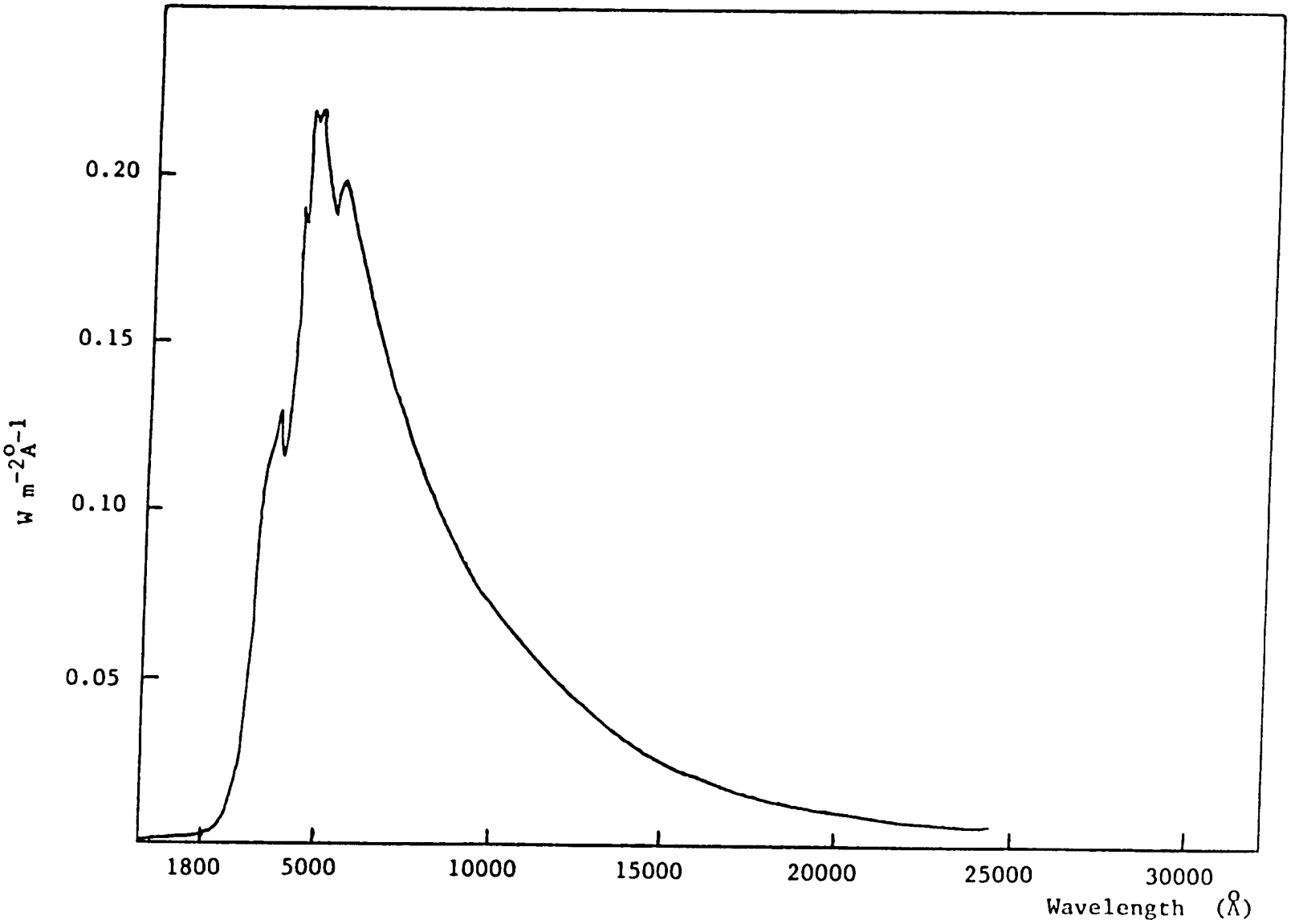


Figure 9. Wavelength dependence of the solar ionizing radiation.

atoms and molecules per unit depth along its path and gives up its ionizing energy as a function of altitude more quickly. Therefore the resulting ionization rate peaks at a higher altitude. In addition to this effect, the neutral atmospheric density is always decreasing with increasing altitude (because of the effect of the force of gravity), causing the rate at which the ions and electrons are lost by chemical recombination to decrease with increasing altitude, thus increasing the number of electrons that remain free. The combination of these two processes gives rise to an altitude peak in the distribution of electrons (see Figure 10). The combination of different solar radiation components and atmospheric constituents results in a multilayered structure, with each exhibiting its own diurnal behavior and with each having its own response to sudden changes in the level of solar activity.

The rate at which the ionosphere responds to diurnal changes in solar radiation (solar zenith angle) is relatively rapid, though the rate varies with the affected layer. At the lower and middle altitudes, the D layer and E layer disappear almost immediately with sunset, while it takes the F_1 layer about an hour to reach nighttime values. In the upper F_2 layer, where transport of ionization becomes an important factor, the time response is slowest, so that it takes some 12 hours for the daytime electron densities to decrease by a factor of 10. The ionosphere reacts in a similar fashion to solar flare activity, responding more rapidly at lower altitudes. These effects are discussed in Section 3.

1.3.2 PARTICLE PRECIPITATION

Particle precipitation is another source of ionization in the earth`s upper atmosphere. Energetic charged particles originating from the sun and trapped in the earth's magnetosphere, spiral toward the earth's atmosphere along magnetic field lines (Figure 11). This process is called precipitation. When the precipitating particles collide with neutral air molecules, they produce free electrons and positive ions. Depending on their initial energy, they give up most of this energy in different altitude regions. For example, 100 eV electrons ionize the atmosphere at F region altitudes (250 km), 1 to 10 keV electrons ionize the E layer (100 km), and greater than 500 keV electrons ionize the D region (60 km).

Even during quiet solar periods, a steady stream of particles approaches the higher latitudes, mostly on the dark side of the earth after being trapped in the earth's magnetic field

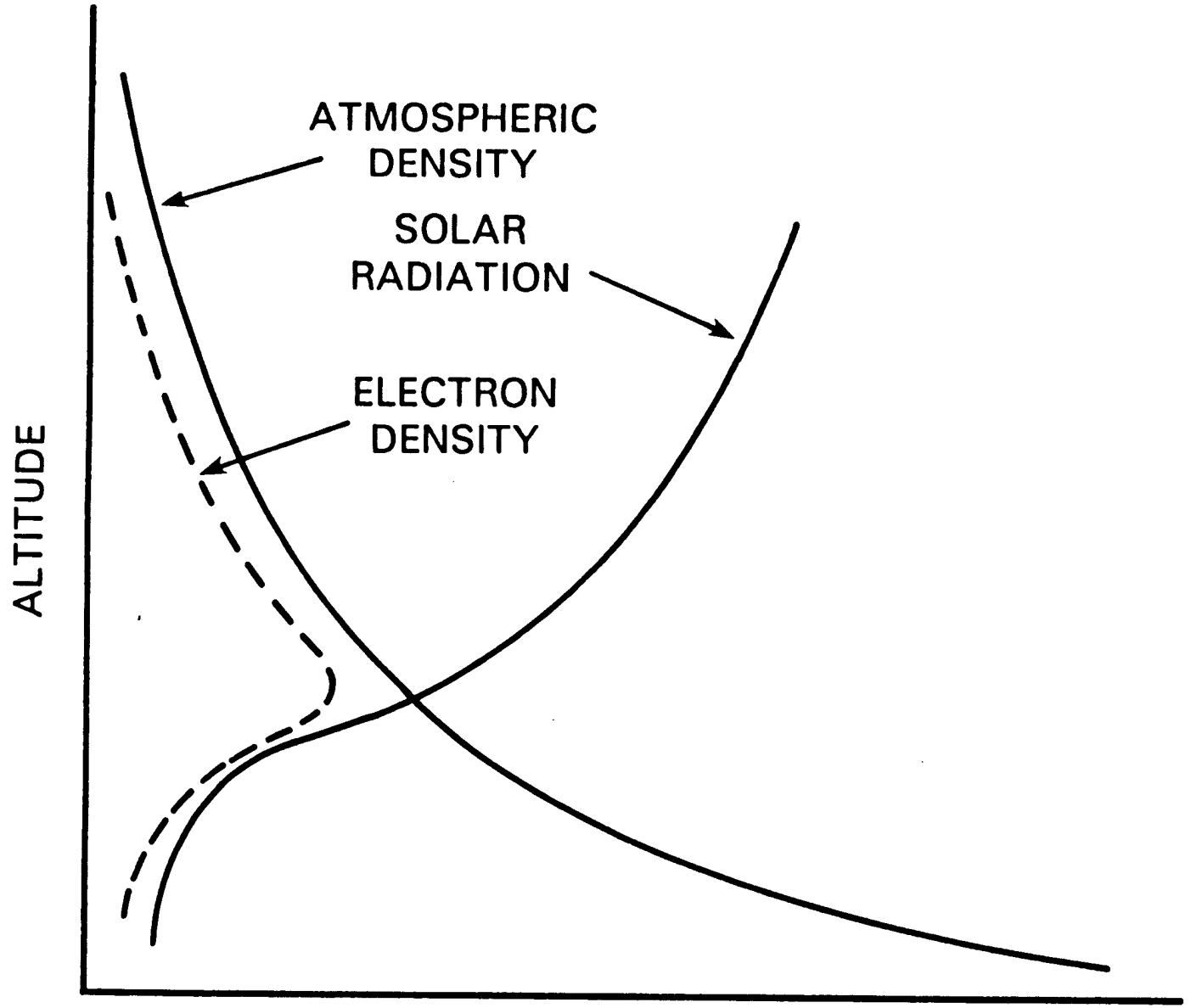


Figure 10. Altitude dependence of atmospheric density, solar radiation and the resulting electron density.

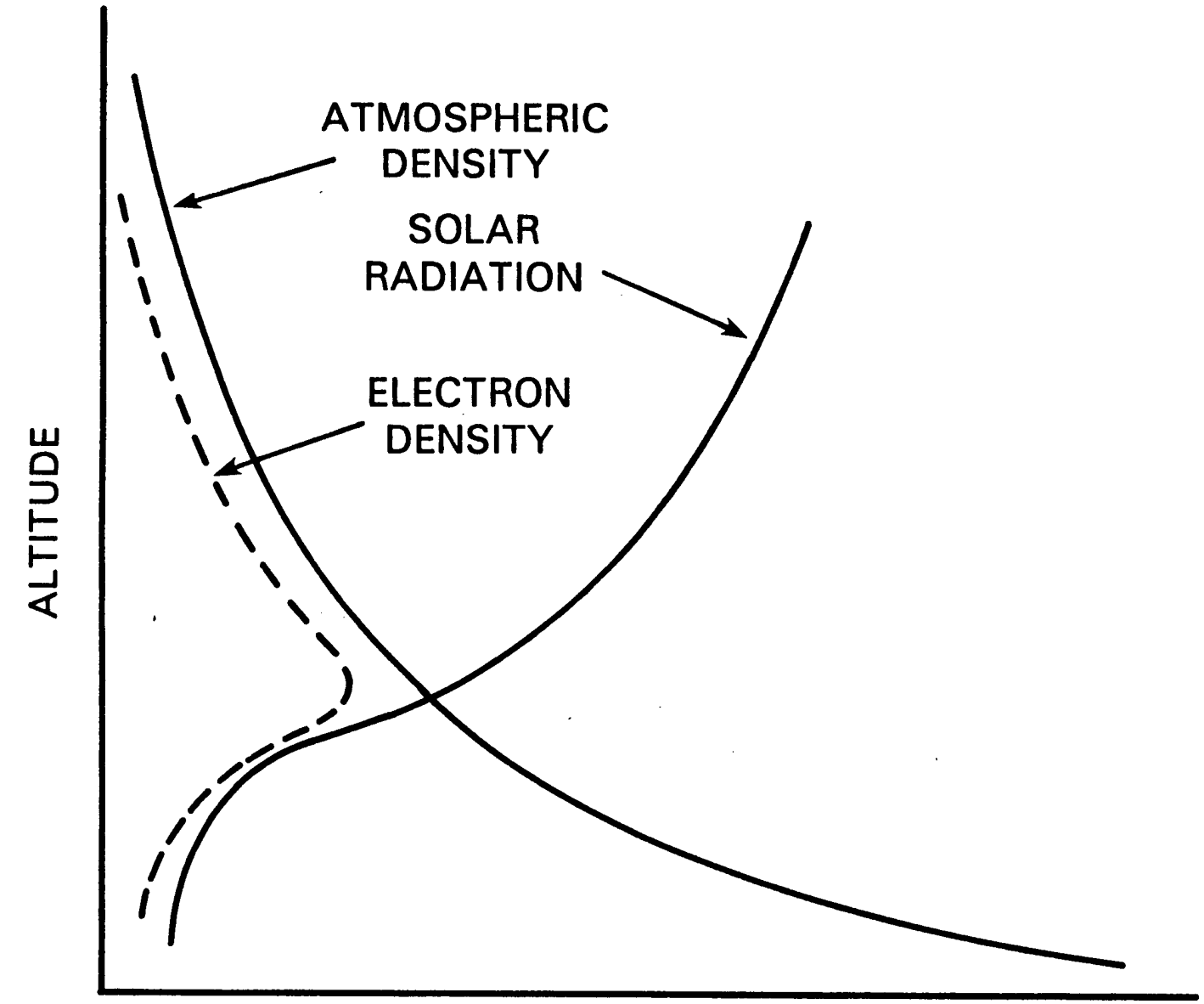


Figure 11. Maintenance of the Ionosphere by particle precipitation.

and stored in the magnetospheric tail. The tail is so named because the interaction of the solar wind with the outer reaches of the earth's magnetic field drags the outer field lines like a tail in the direction away from the sun.

The aurora is a visible manifestation of the collision process between energetic particles and neutral atmospheric molecules and atoms; some of the collision energy of these precipitating particles is converted to visible light, some ionizes the neutrals. The major part of the auroral precipitation is concentrated by the internal structure of the magnetosphere into a band around the geomagnetic pole, the auroral oval. The steady precipitation of particles, which maintains auroral forms along the oval at all times, is intermittently enhanced by intense events known as auroral substorms. The events are due to more intense particle fluxes and often enhanced particle energies, and produce strong enhancements in the high latitude D and E regions. They result in enhanced radar clutter and enhanced losses on the affected OTH propagation paths. The substorms occur at irregular intervals, every few hours, and last typically one to two hours. Substorms become more frequent and intense during geomagnetic storms.

The solar wind is a stream of protons and electrons flowing from the sun at a speed of about 400 km/s. Current theory indicates that these particles originate in the highly ionized corona of the sun, carrying the solar/interplanetary magnetic field lines towards the earth, as well as radially away from the sun in other directions that do not directly influence the earth. The solar wind particles travel for about four days before arriving at the sunward boundary of earth's magnetic field, where they interact with it. These solar wind particles are able to enter the earth's magnetic field envelope in the tail region of the magnetosphere. From there, through complex processes, they are further energized, and they travel back towards the earth along the earth's field lines and produce a more or less steady precipitation of ionizing particles into the dark side of both the northern and southern high latitude regions. On the day side, solar wind particles find direct access into the atmosphere, precipitating into the so called cusp or cleft region (for details see later in this section). Suffice it here to state that the topography of the cleft and night-time precipitation regions overlap in the morning and evening sectors, thereby forming the auroral oval. When the sun is disturbed (active), the solar wind is enhanced and, after some time, allowing for the transit of the particles from the sun to the earth (under disturbed conditions this

typically takes 24 to 48 hours) produces enhanced particle precipitation and increased ionization at high latitudes. The supply of precipitating particles is constantly replenished by the solar wind.

2. SOLAR AND GEOMAGNETIC ACTIVITY

2.1 Introduction

Solar electromagnetic radiation generates and maintains the global ionosphere. Solar disturbances, known as solar flares, which are often accompanied by intense X-ray bursts, have a strong impact on the ionosphere and on radio propagation. Besides electromagnetic radiation, the sun also emits a steady stream of electrons and protons, the solar wind. The solar wind carries with it field lines of the solar magnetic field. A few solar radii away from the sun, this extended magnetic field is called the Interplanetary Magnetic Field (IMF). The flux of solar wind particles that can enter into the magnetic field structure of the earth depends critically on the configuration of the IMF in the vicinity of the earth. The solar wind interacts with the earth's magnetic field, deforming it in the process into the magnetosphere. Through rather complicated processes, some of the flux of solar wind particles precipitate into the high latitude ionosphere. The internal structure of the magnetosphere guides these particles into two bands encircling the poles (the northern and southern auroral ovals and adjacent regions), and under quiet conditions also into the polar caps, the regions inside the auroral ovals.

Solar flares are accompanied by localized enhanced ejection of solar plasma, primarily electrons and protons. These particles form large plasma clouds, which travel from the sun through the interplanetary space. If such a plasma cloud reaches the earth, it interacts with the earth's magnetic field, generating magnetic storms of varying magnitude, and enhanced particle precipitation into the ionosphere. Such precipitation occurs in an enlarged auroral oval, generally in the form of auroral substorms. Auroral substorms are accompanied by bright and active auroral displays, and result in strong auroral currents (auroral electrojet). Both the particle precipitation and the auroral electrojet cause intense ionospheric electron density irregularities, a well known source of Doppler spread clutter for the OTH radars operating in the vicinity of the auroral latitudes. With enhanced activity, the oval increases in diameter, causing enhanced auroral disturbance effects to occur deeper in the OTH coverage areas. For details on auroral effects on OTH, see Section 3.2.1 (on the High Latitude Ionosphere). Mainly, the ECRS (East

Coast Radar System) segment 1 and WCRS (West Coast Radar System) segment 3 (both north looking sectors) are affected by the high latitude auroral phenomena. But due to the range folding property of the OTH radar system the long/far range equatorial clutter when present can affect these north looking sectors.

Of more large scale impact on the OTH radars however, are magnetic storms. The increased energy input into the atmosphere during these storms heats the atmosphere and substantially enhances the neutral density at ionospheric heights. This in turn increases the recombination of ionospheric plasma, often drastically reducing the electron density (therefore critical frequency) in the ionosphere, forcing the OTH radars to operate at low frequencies, competing with other users of the HF spectrum for the remaining small usable frequency band. Enhanced absorption and electron density irregularities generating auroral clutter also accompany the magnetic storms, further complicating the operation of the radars.

Because of the importance of the sun to the generation of the ionosphere, and because of the drastic effects solar activity has on the ionosphere and on systems operating through it, ionospheric physicists have for a long time studied the relationships between the quiet and the disturbed sun and the ionosphere.

To allow the radar operator to understand the complexity of the ionosphere, which he should consider as part of the OTH system, this section provides some insight into the solar phenomena which control the ionosphere, the quantities used to describe the state of the sun, and the resulting geomagnetic activity. Solar and geomagnetic characteristics are discussed to provide an understanding of the cyclic nature of 1) the solar behavior (the sunspot cycle), of 2) the occurrence patterns and durations of major storms, and of 3) the recurrent nature of certain types of magnetic activity. Various solar and geomagnetic characteristics and their statistical behavior, and where relevant, their relations to ionospheric conditions are discussed.

2.2 Solar and Magnetic Activity Indices

Sunspots have been observed for hundreds of years. One of the first objects that Galileo (Chapter XXVI, Chapman and Bartels[5]) studied with his first tiny telescope was the sun, which he observed by projecting an image of the sun onto a white screen, and what he saw were dark spots, the sunspots.

Sunspots have been observed and recorded systematically for nearly 400 years. Single spots are relatively rare; usually spots tend to occur in compact groups. The Swiss astronomer Wolf (see Haymes[6]) designed an index of solar activity and used it to create a reasonably homogeneous time series going back to the eighteenth century with monthly averages, and to 1818 with daily averages. His index, now called R_z is defined by the equation:

$$R_z = f(10 x + y)$$

where f is a scale factor near unity for a given observatory, x is the number of sunspot groups, and y is the total number of spots.

We have daily records of sunspots (with some gaps) back to 1818 and a continuous record beginning in 1849. Figure 12 shows the daily sunspot counts (R_z) versus time since 1818. Notice the gaps appearing for R_z less than 10. Even one group of spots, visible on the disc, forces R_z to jump from zero to ten. Notice also the highest daily sunspot count ever observed, 355, on December 24 and 25, 1957, visible on Figure 12 just before 1960.

Monthly means of R_z are available even further back in time to 1749 as seen in Figure 13 which shows the monthly average sunspot number, R_s, versus year for 250 years. There are quasi-periodicities evident in the complete record; the most obvious, more or less regular solar rhythm is called simply the "eleven year sunspot cycle". So dominant is the appearance of the 11-year cycle that the individual cycles have been assigned numbers beginning with cycle 1 in 1755 to cycle 22, the current cycle, which began in August 1986, and reached its maximum in July 1989, with a monthly mean of 127, and a smoothed monthly mean sunspot number of 158. However, the time series of monthly average sunspot numbers is in no sense stationary, not even close! The period of the prominent 11-year cycle, measured between successive minima, actually varies from less than ten years to almost twelve years. The amplitude is also highly variable, with maximum counts varying by a factor of more than four to one. The time from a minimum to the next maximum varies from about 2 1/2 to 6 years.

Since the sunspot time series appears so bumpy and noisy, it has become customary to use a simple smoothing technique to remove many of the high frequency jiggles. One of the smoothing techniques is to determine the average from two successive 12-month running means (compute 12 month running mean once, and compute a second 12 month running mean from this new computed. data set). Using this technique results in the curves in Figure 14, which shows

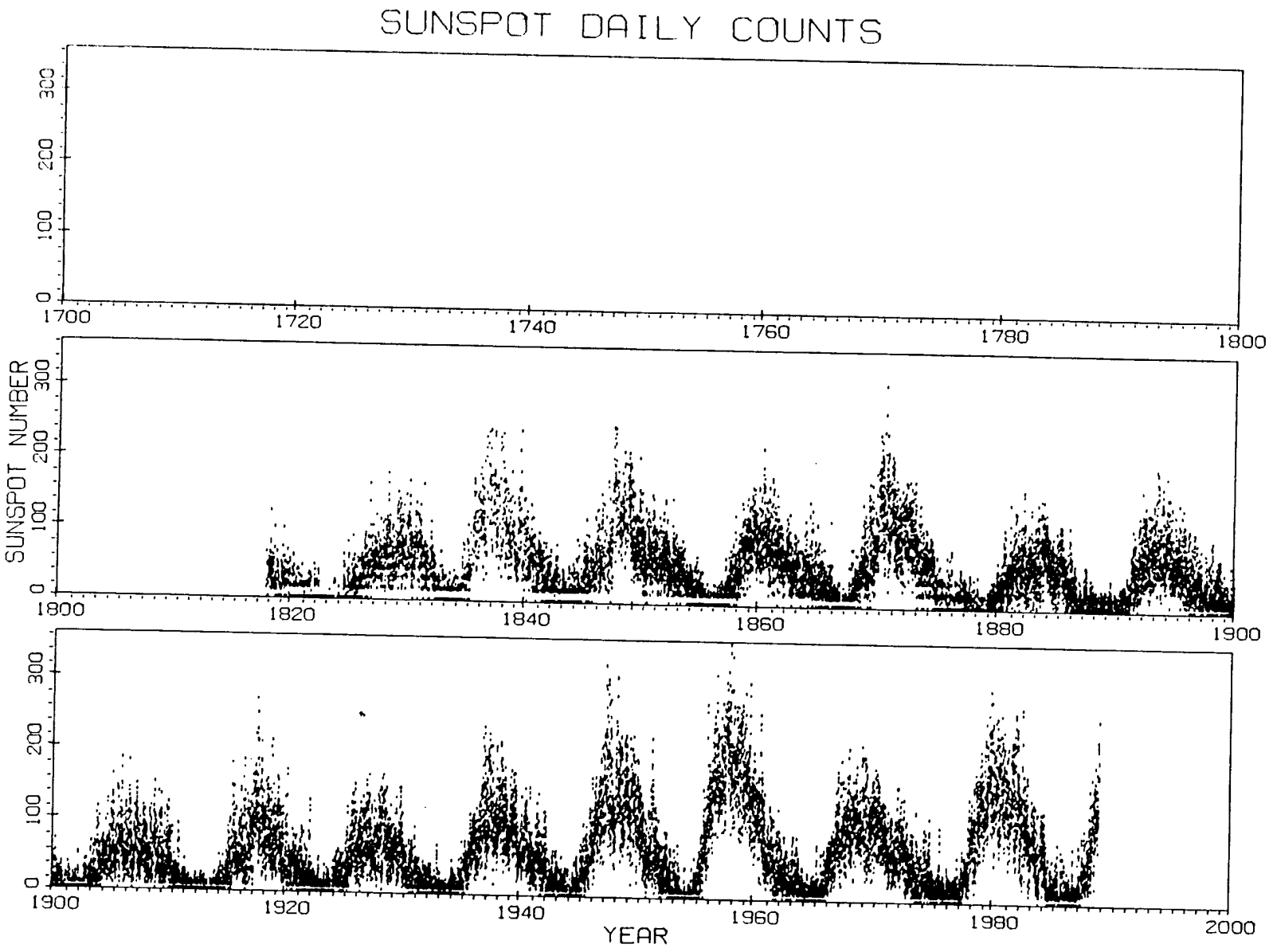


Figure 12. Daily sunspot number for years 1816 to 1990.

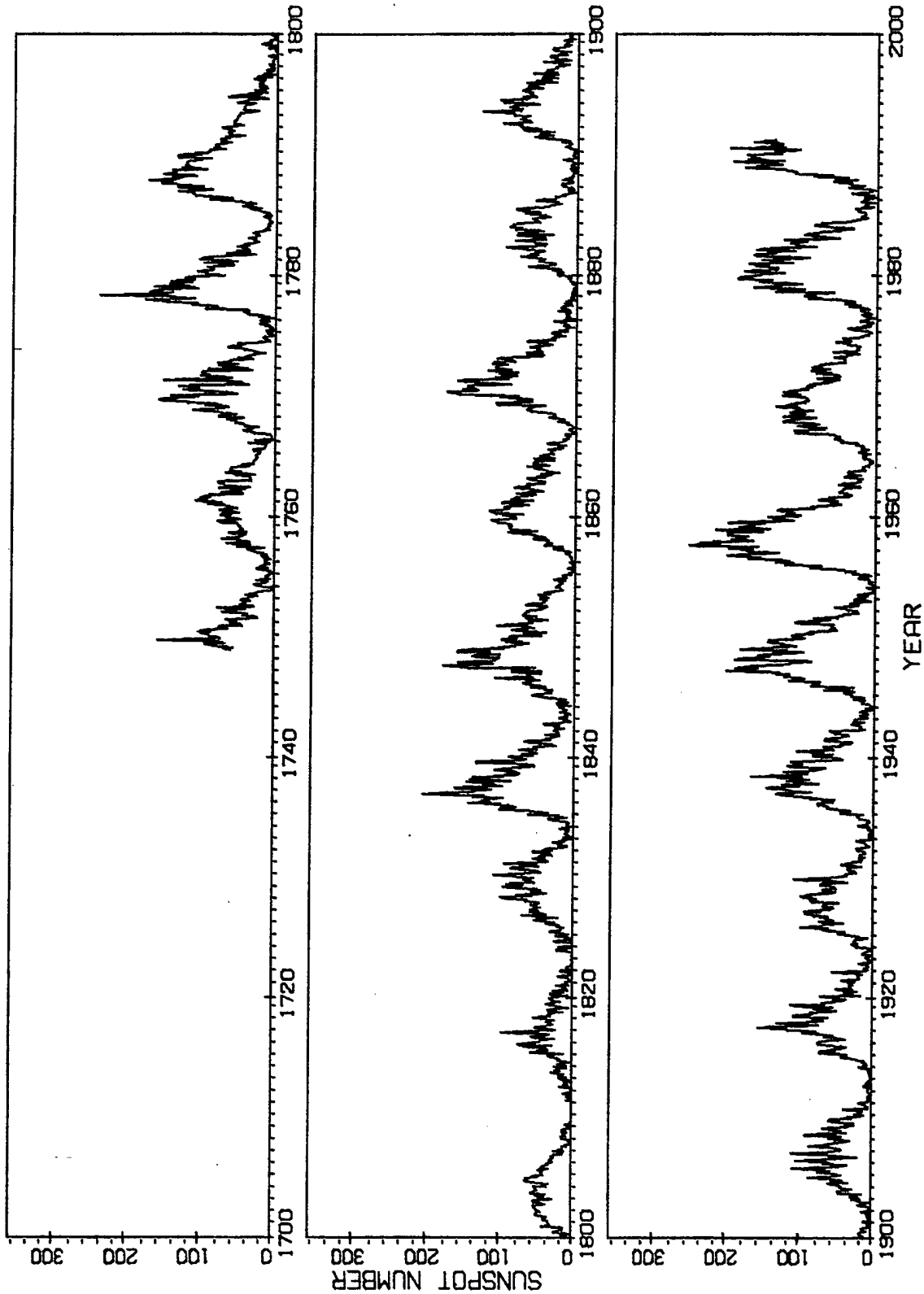


Figure 13. Monthly mean sunspots for years 1750 to 1990.

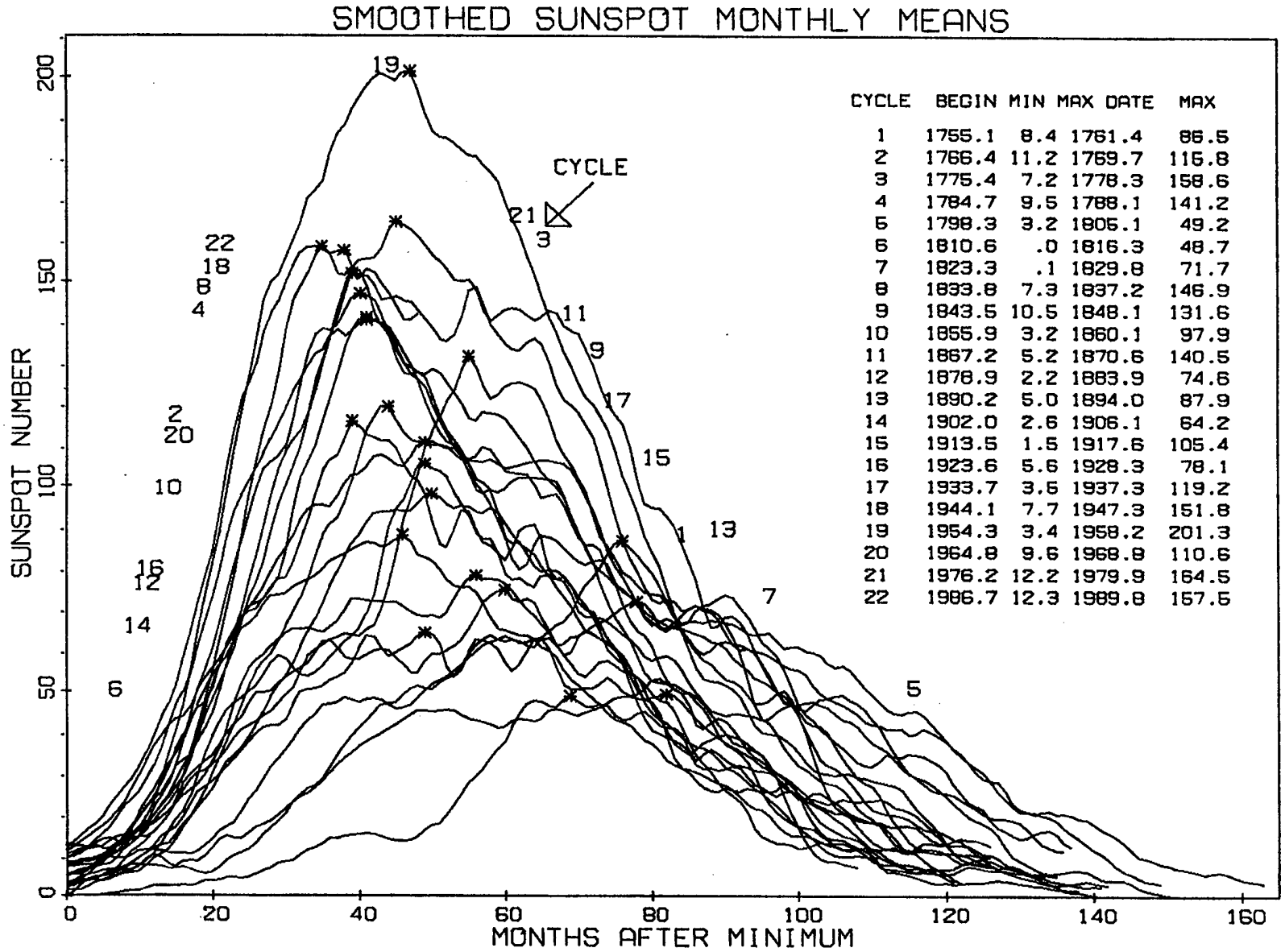
the shape of each solar cycle with an asterisk plotted at the maximum of these smoothed numbers. An inset table in the figure displays the cycle number, the beginning date (conventionally set at the smoothed minimum) the sunspot count at minimum and the date and magnitude of maximum. It can easily be seen that the highest maxima occur after a large quick rise and that the size of the maximum is inversely proportional (at least approximately so) to the time between a minimum and the next maximum.

Figure 15 shows the current status of cycle 22 as well as the average of cycles 1-21, and the average plus and minus the standard deviation. The maximum and minimum sunspot numbers ever observed are also shown in the figure.

Longer periods, notably about 80-years, while not clearly evident in the Figures, have been found in a remarkable, thousand year long record of Chinese auroral observations. For the Chinese records the auroral counts act as proxies for the solar activity.

On shorter time scales the rotation period of the sun is quite important geophysically. In fact the rotation of the sun was observed and quantified by watching sunspots. The sun, however, does not rotate as a solid body; in fact the period varies from less than 24 days at the equator to more than 34 days at higher solar latitudes. Taking into account the earth's rotation, the effective period of about 27 days is often used in solar-terrestrial calculations and data reduction. (The geomagnetic storms originating in certain sunspots do recur with the recurrence of these sunspots from solar rotation). The rotations have been assigned numbers, called "Carrington Rotation Numbers", starting with number 1 in 1848.

Figure 16 shows the frequency distribution of daily sunspot number. This histogram is plotted with an interval of 10 sunspots. For 20 percent of the time, the sun is quiet, showing less than 10 spots. The bin from 0 to 10 spots contains contributions only from individual sunspots (no sunspot groups), because if even one spot group is visible the count must jump to 10. The hole (very quiet sun with zero- no sunspots) is evident in Figure 12 (for example, the bold dark line along the horizontal axis in the bottom left hand corner of the figure for the years 1901-1903). The mean of all 59211 data points is about 52. The median value, 58, indicates that there are as many days with counts less than 58 as there are days with counts greater than or equal to 58. Of course the counting statistics change markedly with position in the 11-year cycle. For example, the zero spot case occurs only close to sunspot minimum.



CYCLE	BEGIN	MIN	MAX DATE	MAX
1	1755.1	8.4	1761.4	86.5
2	1766.4	11.2	1769.7	115.8
3	1775.4	7.2	1778.3	158.6
4	1784.7	9.5	1788.1	141.2
5	1798.3	3.2	1805.1	49.2
6	1810.6	.0	1816.3	48.7
7	1823.3	.1	1829.8	71.7
8	1833.8	7.3	1837.2	146.9
9	1843.5	10.5	1848.1	131.6
10	1855.9	3.2	1860.1	97.9
11	1867.2	5.2	1870.6	140.5
12	1878.9	2.2	1883.9	74.6
13	1890.2	5.0	1894.0	87.9
14	1902.0	2.6	1906.1	64.2
15	1913.5	1.5	1917.6	105.4
16	1923.6	5.6	1928.3	78.1
17	1933.7	3.5	1937.3	119.2
18	1944.1	7.7	1947.3	151.8
19	1954.3	3.4	1958.2	201.3
20	1964.8	9.6	1968.8	110.6
21	1976.2	12.2	1979.9	164.5
22	1986.7	12.3	1989.8	157.6

Figure 14. Monthly mean smoothed sunspot number for individual solar cycles.

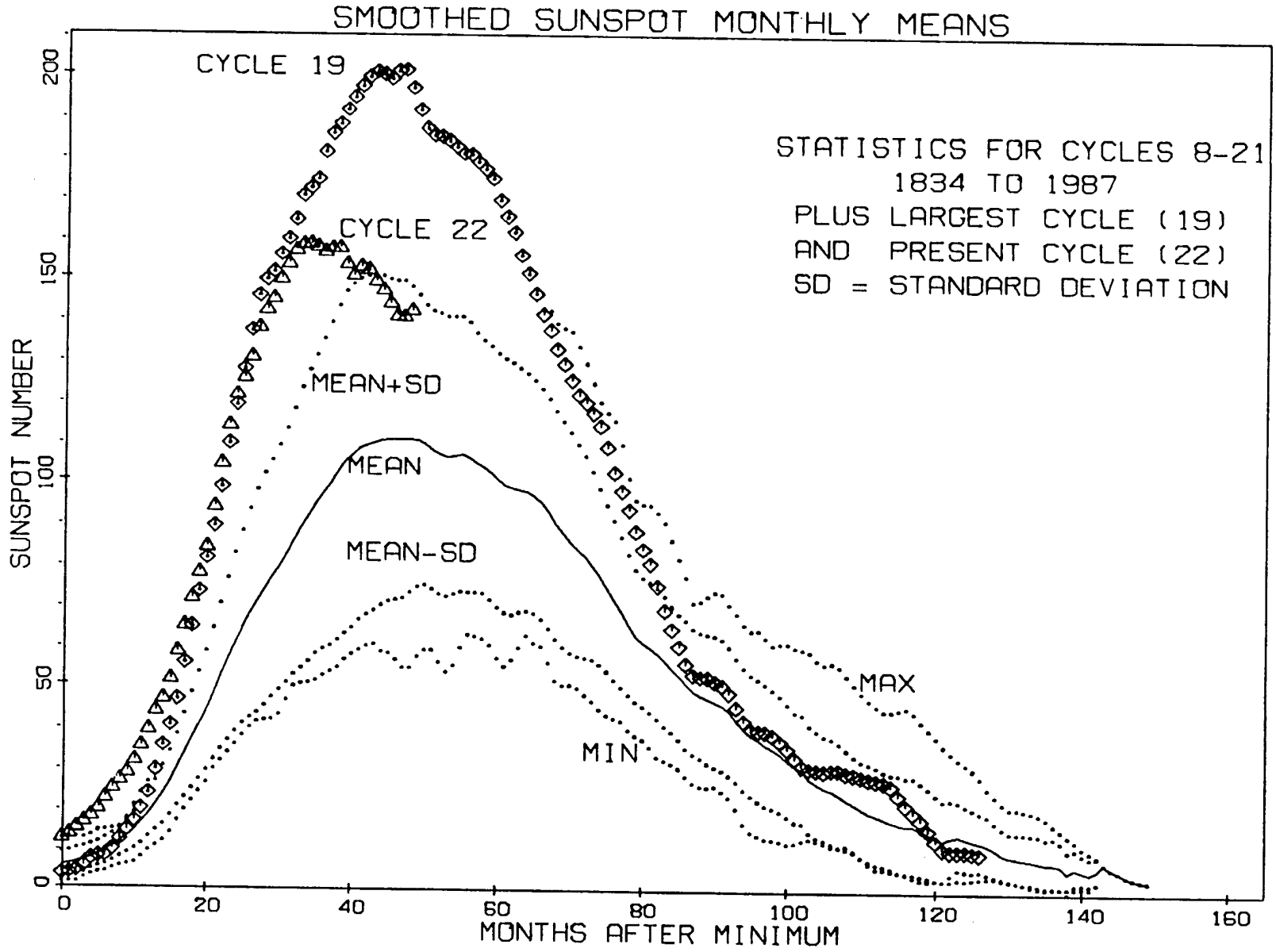


Figure 15. Comparison of sunspot cycles 19 and 21 with maximum mean and minimum sunspot cycles.

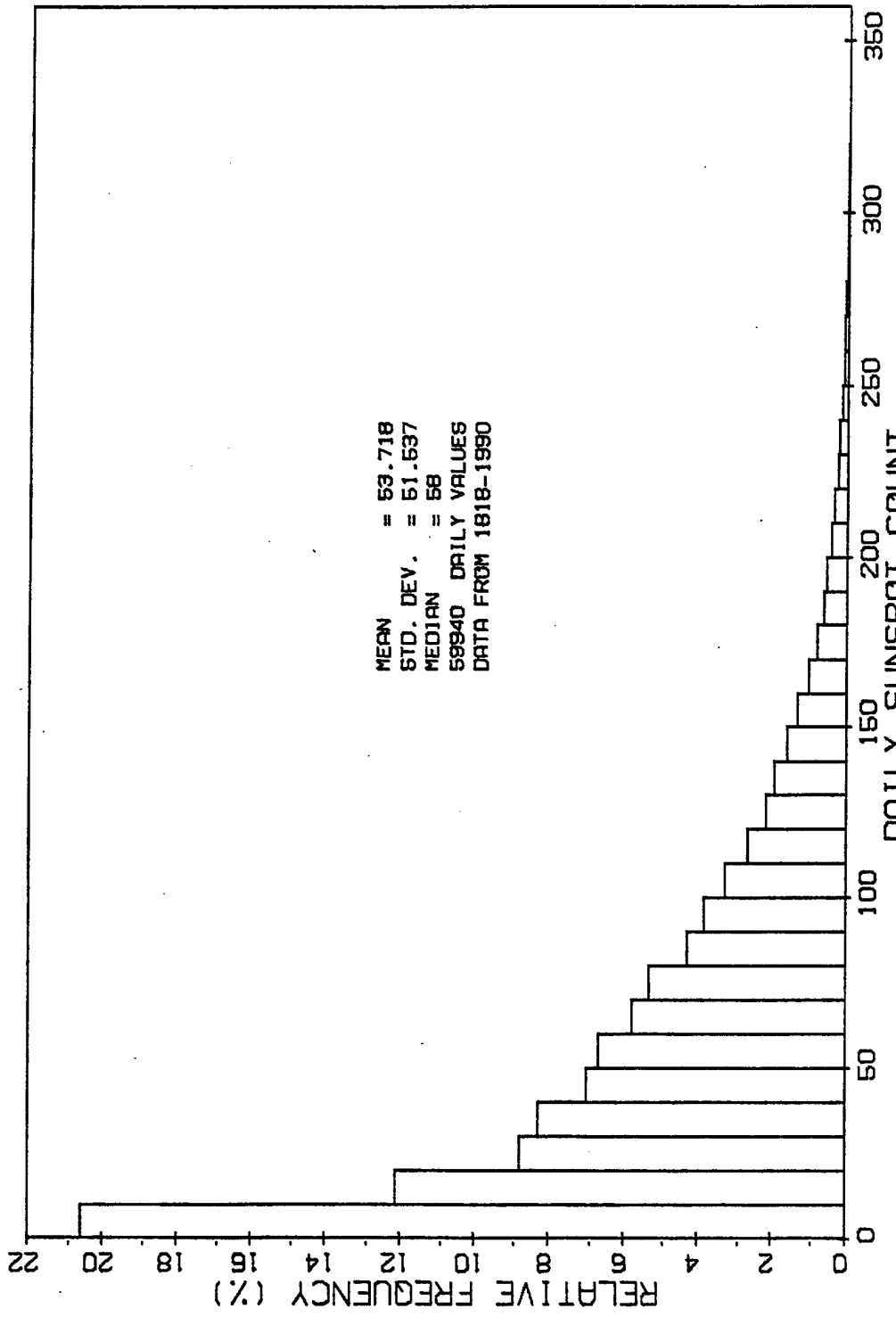


Figure 16. Relative occurrence of daily sunspot number.

The sun is also a very powerful emitter of electromagnetic radiation in frequency bands in addition to the visible. This radio flux, which originates from atmospheric layers high in the sun's chromosphere and low in the corona, changes gradually from day to day, in response to the number of sunspot groups on the disk. Radio intensity levels consist of emissions from three sources: from the undisturbed solar surface, from developing active regions, and from short-lived enhancements above the daily level. Solar flux density at 2800 MHz (10.7 cm) has been recorded routinely by radio telescope near Ottawa, Canada since Feb. 14, 1947. Each day, levels are determined at local noon (1700 GMT) and then corrected to within a few percent for factors such as antenna gain, atmospheric absorption, bursts in progress, and background sky temperature.

The significance of the sunspot number for the ionosphere is borne out by the fact that a high correlation exists between the smoothed mean monthly mean sunspot numbers and the monthly medians of the hourly f_0F_2 (maximum F-layer density) values. The CCIR coefficient set describing the monthly median hourly global f_0F_2 maps uses the smoothed monthly sunspot number as the controlling parameter. The relationship between these f_0F_2 medians and the smoothed sunspot numbers is approximately linear as is evident in Figure 6.

The 10.7 cm solar radio flux is so highly correlated with the sunspot number R_z that either variable may be transformed to the other, with little loss of information. Figure 17 shows smoothed monthly means of sunspot number versus solar flux at 10.7 cm. The best fitting (least squared deviation) straight line and its equation are also shown. The curve shows a definite bend in the vicinity of R_z less than about 20. It should be noted that the 10.7 cm solar flux does not ionize the atmosphere; the flux is, however, highly correlated with the EUV, which does ionize the atmosphere.

As might be expected, in addition to the rather regular behavior described so far, the sun produces very large outbursts of energy across the entire spectrum of solar emission. Indeed, when the sunspot count is very high, at the time called solar maximum, giant white light flares are common. These flares may produce large fluxes of solar protons and electrons that travel as interplanetary shock waves at high speed, reaching the earth's outer magnetosphere within 24 to 48 hours, often producing "sudden-commencement" magnetic storms. When a shock wave hits and compresses the magnetopause boundary, the earth's magnetic field jumps by 10 to 20 nT (see Figure 18) or so, nearly simultaneously all over the earth. The accompanying magnetic storms,

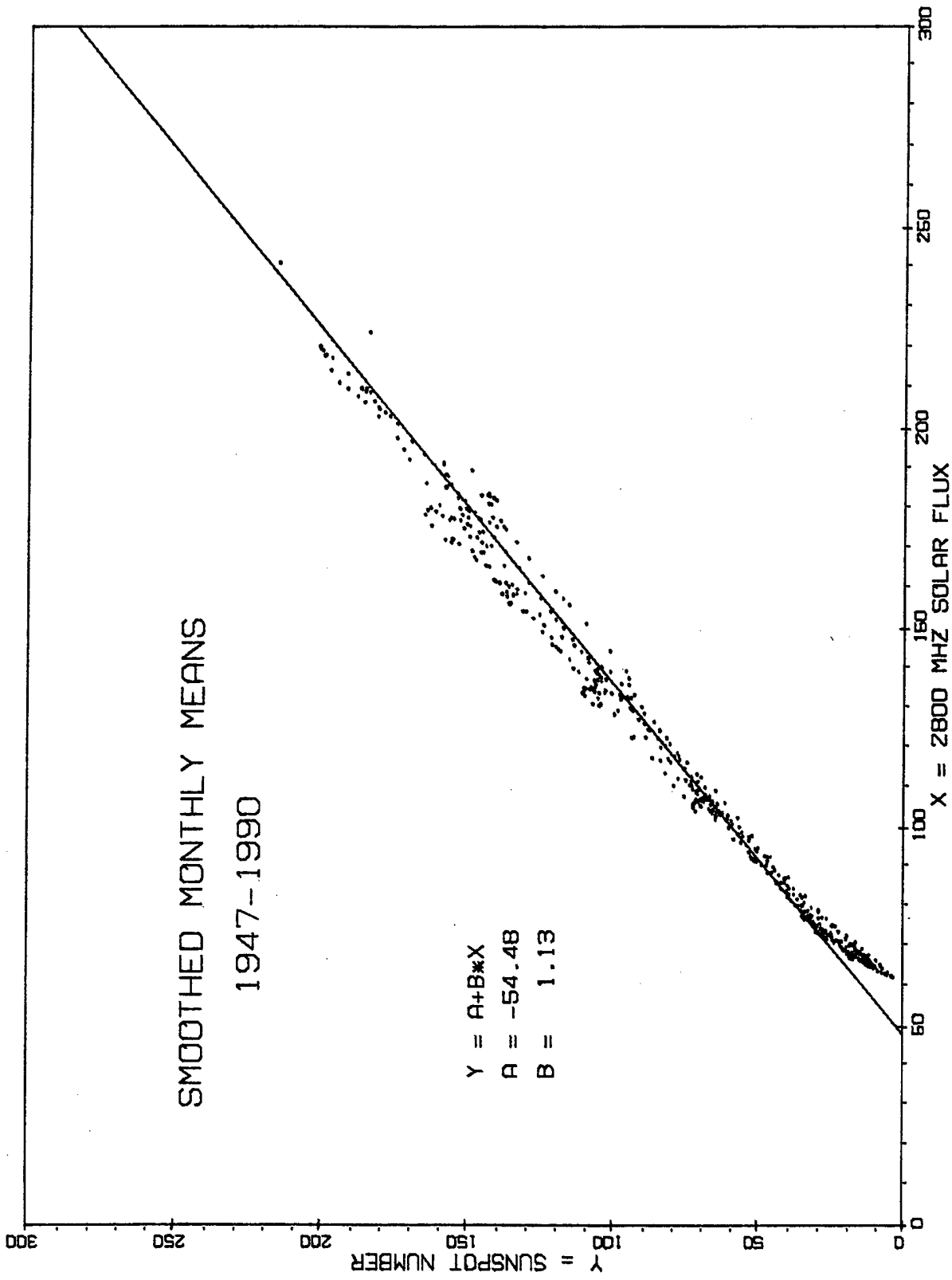


Figure 17. Relationship between sunspot number and 2800 MHz (10.7 cm) solar flux.

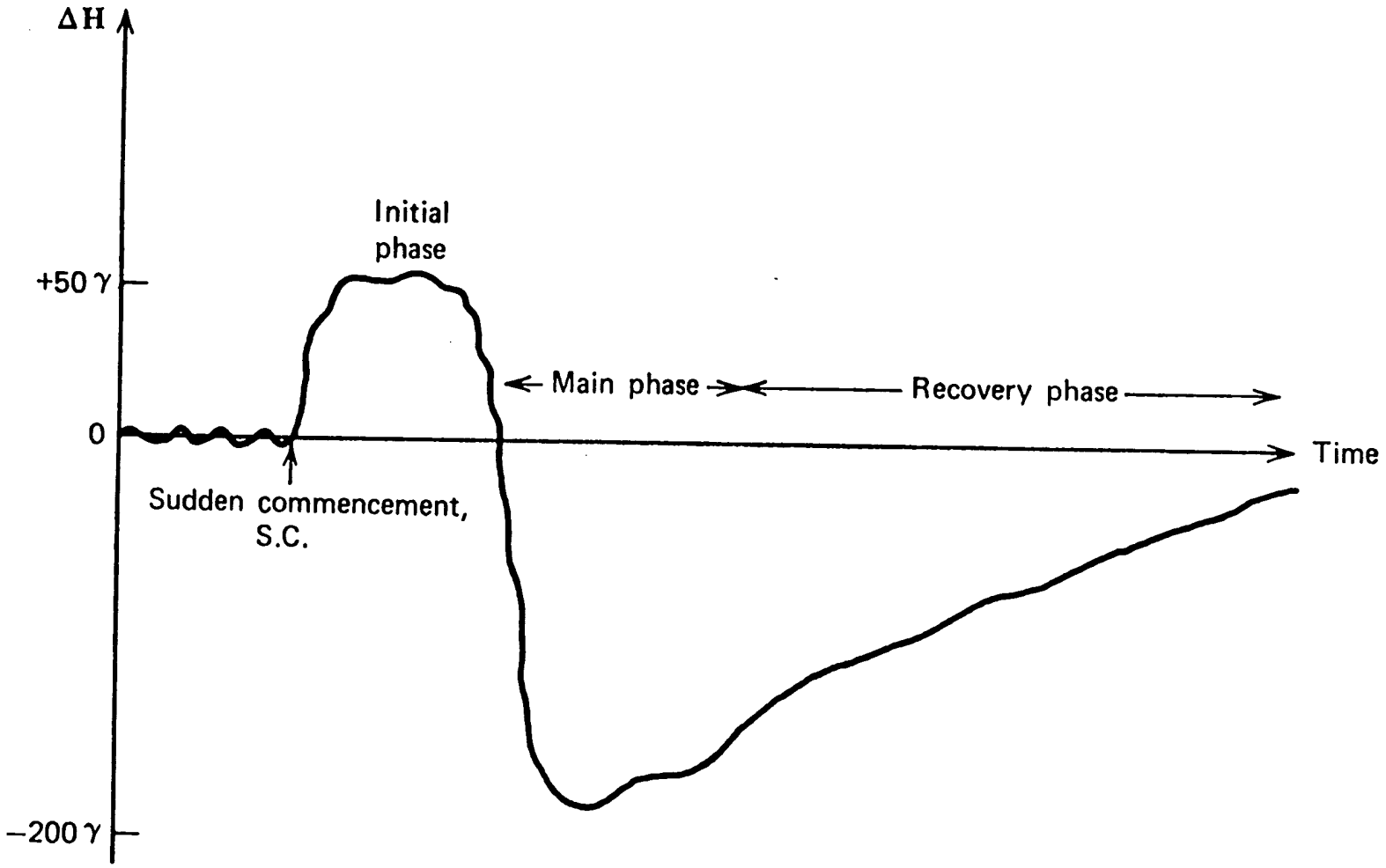


Figure 18. Schematic of a typical midlatitude geomagnetic storm, (Haymes[6], p.228).

lasting from one to several days, can cause severe damage to electrical power grids in auroral zone areas and knock out communication systems for extended periods of time through their effects on the ionosphere.

At other times in the solar cycle, typically during the descending portion just after solar maximum, long lived regions, originally called "M regions", cause recurrent magnetic storms on earth. The period of recurrence is about 27 days, the average solar rotation period. It is now known that the "M regions" are coronal holes, which allow the sun's magnetic field easy access to interplanetary space and easy coupling to the earth's magnetosphere.

2.3 Solar Wind and Magnetosphere

The main magnetic field of the earth, a vector with three components: X (north), Y (east) and Z (vertical down), (see Figure 19) is almost entirely of internal origin. The magnitude of the field near the equator is about 30,000 nT. Near the poles the magnitude is about 50,000 nT. The magnitude of the horizontal component is called H, and the angle in the horizontal plane between the horizontal component and true north is called the magnetic declination, D. The angle by which a magnet's needle dips below the horizontal plane is called the Dip or inclination, I. The earth's magnetic field can be approximated as a simple dipole centered at the earth's center; and is very well represented by an eccentric dipole shifted from the earth's center, (See Figure 20). Finer details can be recovered with a full expansion of the geomagnetic field expressed in spherical harmonics. The field arises from a self- sustaining dynamo in the earth's molten (mostly nickel and iron) interior. This internal field is nearly constant with only a small, very slowly varying secular variation. Relatively small regions of magnetic anomalies due to metallic ore deposits are also common.

The dipole field of the earth would extend to great distances if there were no external influences upon it, but the sun is constantly emitting a hot "solar wind" plasma consisting mostly of hydrogen and helium nuclei (protons and alpha particles) and electrons. The solar wind travels in a generally radial outward direction from the rotating sun and carries embedded within it some of the sun's magnetic field. The solar rotation produces a "garden-hose" effect (see Figure 21A) in which the plasma flow as seen by an observer sitting above the sun's field is predominantly either toward or away from the sun in an even number of very large sectors (see Figure 21B).

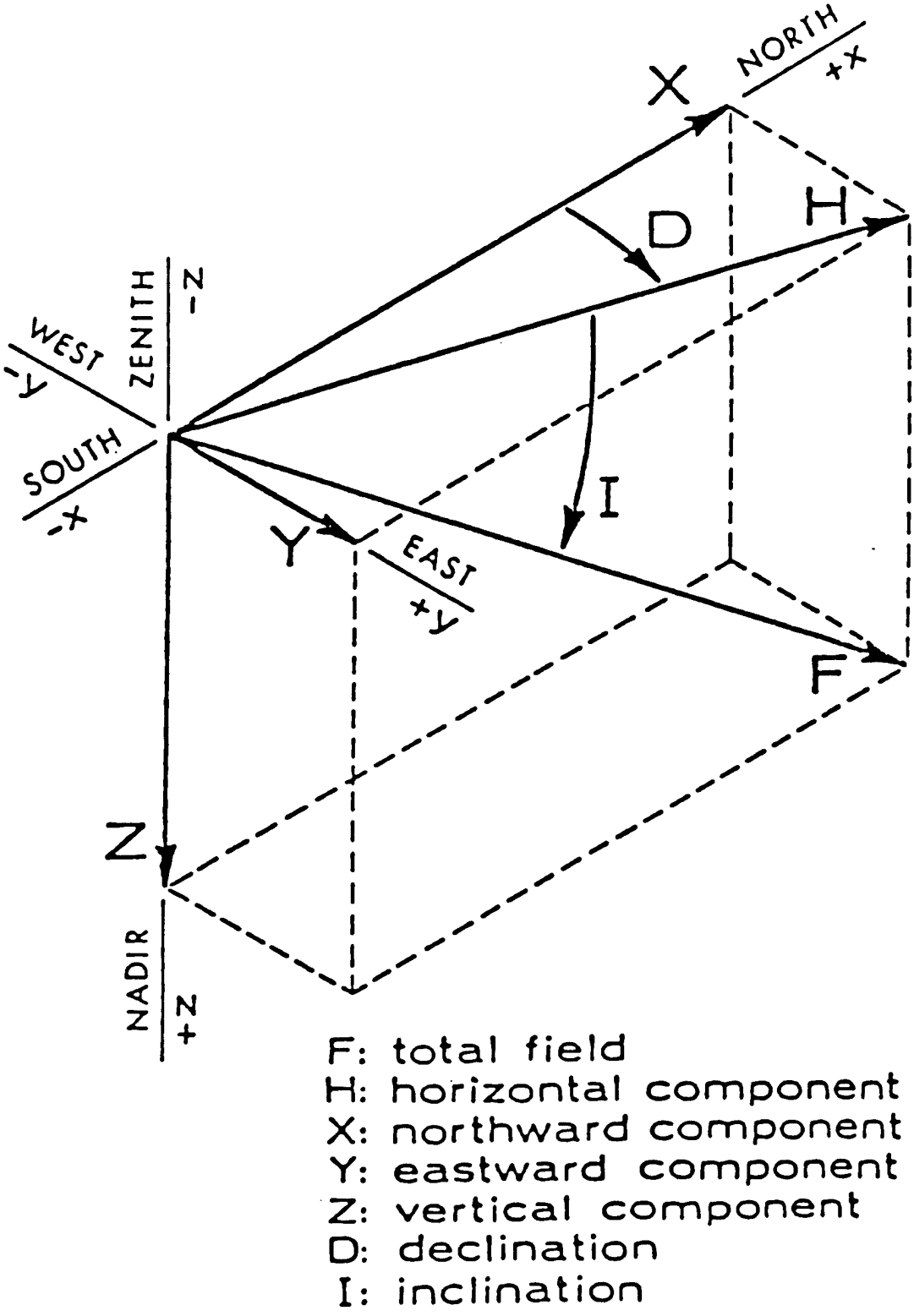


Figure 19. Definition and sign convention for the elements of earth's magnetic field, (Jursa[1], p.4-2).

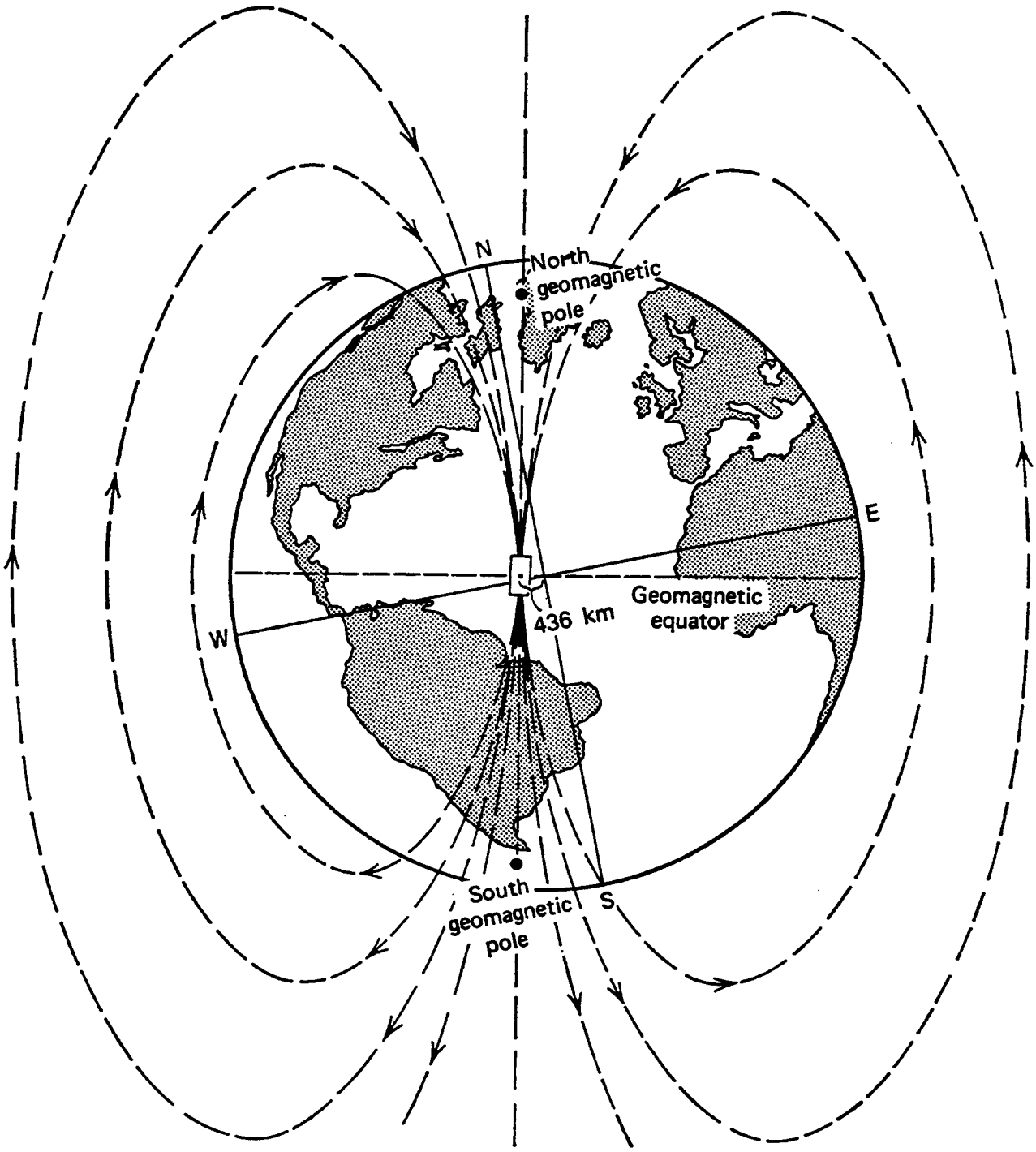


Figure 20. The eccentric-dipole model of the earth's magnetic field(schematic view). The equivalent dipole is ~436 km distant from the center of the planet and is closest to the surface in the hemisphere that contains the Pacific. Hence at a given altitude the field is stronger over the Pacific than it is over the Atlantic. The geomagnetic axis is tilted ~ 11.5° with respect to the earth's rotational axis (the N-S line in the figure), (Haymes[6], p.215).

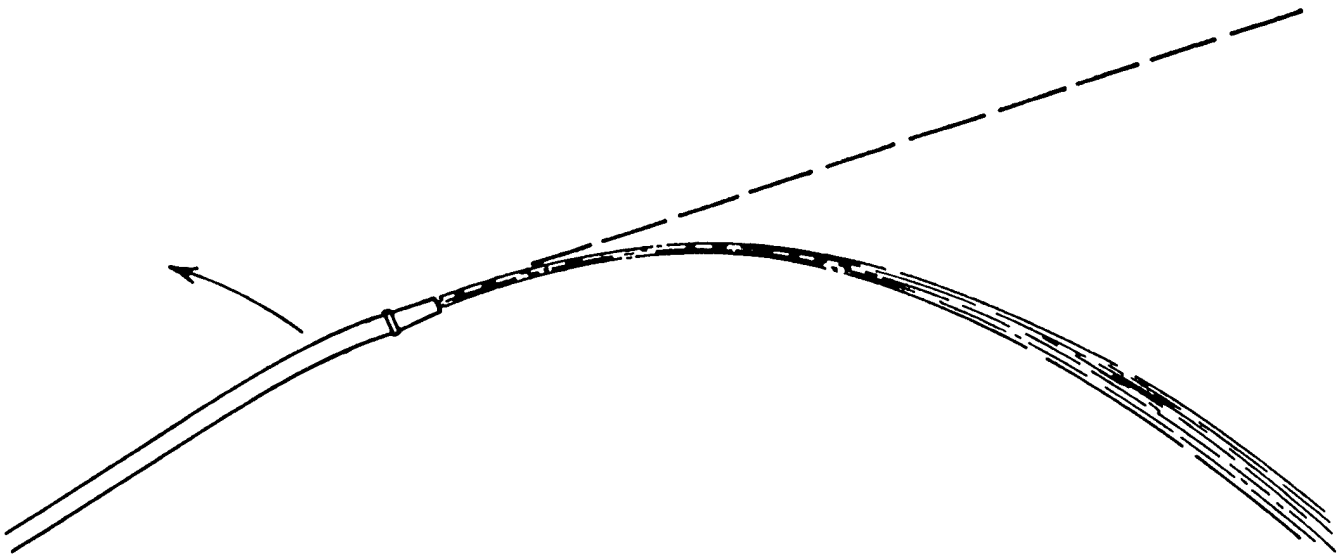


Figure 21A. The spiral path (garden hose effect) exhibited by water issuing from a moving hose, (Haymes[6], p.283).

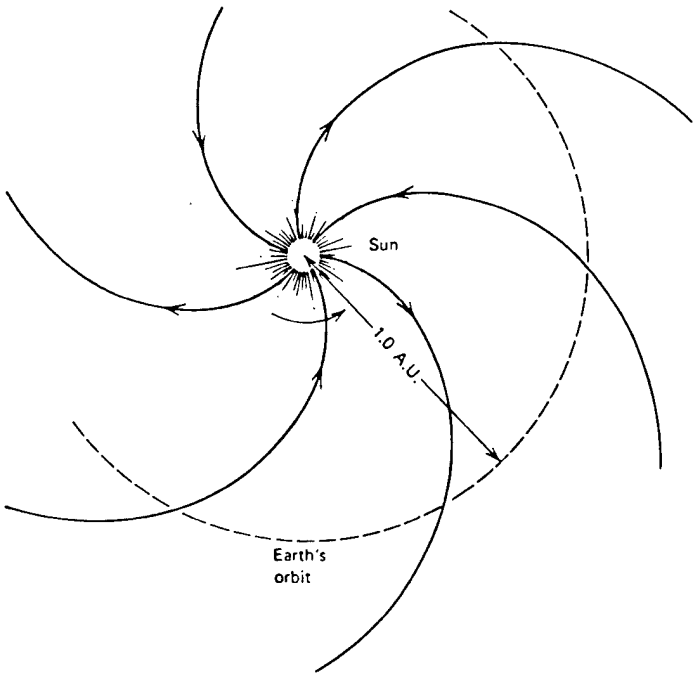


Figure 21B. Schematic representation of the Archimedes spiral structure of the interplanetary magnetic field in the ecliptic plane, (Haymes[6], p.284).

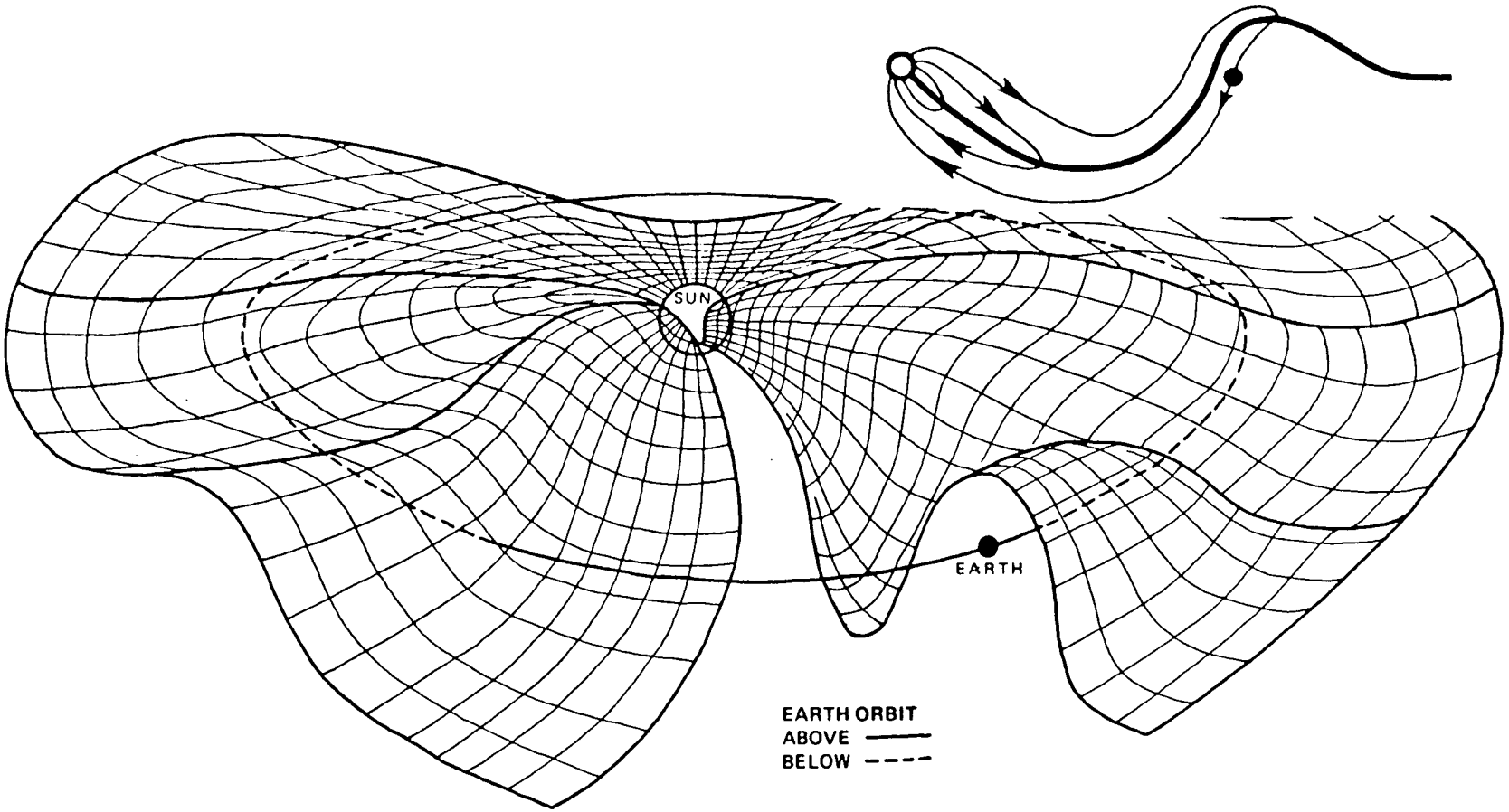
The away (toward) field lines originate above the solar magnetic equator and the toward (away) field lines originate below the equator. During the time of sunspot minimum, the solar magnetic field changes polarity, with "north" poles becoming "south" and vice versa. What had been the "away" hemisphere becomes the "toward" hemisphere and vice-versa. This results from the wavy nature of the location of the solar magnetic equator which the solar wind projects out into space as a wavy "neutral sheet". The earth in its orbit moves sequentially from below the neutral sheet to above it and back, and thus experiences the change from toward to away sectors and back (See Figure 22).

As shown schematically by Figure 23, the effect of this hot, supersonic magnetized plasma, the solar wind, is to confine and compress the earth's magnetic field on the sunward side, and to create an extremely long tail of lowered plasma density and speed on the opposite side. Magnetic field lines originating in one hemisphere and ending in the other are called "closed" field lines. Field lines originating at high north and south latitudes and which extend away from the earth, are called "open". The demarcation line between open and closed field lines defines the polar cap. The auroral oval is found just equatorward of the polar cap region. The funnel shaped region between the dayside closed field line region and the first open field line is exposed to direct access of solar wind plasma.

This region is called the cusp or noon sector of the auroral oval. Imbedded in the magnetotail, the last closed magnetic field line in the night hemisphere is within the night sector of the auroral oval. In this region, energetic particles precipitate out of the tail region, generating auroral substorms, which have strong ionospheric effects on the OTH radars operating in the vicinity.

2.4 Terrestrial Magnetic Activity

In addition to the permanent geomagnetic field, the very slow secular variation, and local magnetic anomalies, all of which are of internal origin, there are many distinct kinds of magnetic variations with time scales varying from minutes to a few days caused by external forces whose origin is on the sun. Days which are truly quiet magnetically are quite rare, occurring less than 3 percent of the time. On such days there are still two reasonably regular variations, called Sq, for solar quiet, and L for lunar. The magnitude of the change in the horizontal component of the geomagnetic field is on the order of 20 gammas for Sq and an order of magnitude less for L (a



THE SOLAR CURRENT SHEET

Figure 22. Three-dimensional sketch of the solar equatorial current sheet and associated magnetic field lines. The current sheet is shown as lying near the solar equator with spiraled, outward pointing magnetic fields lying above it and inward pointing fields lying below it. The current sheet contains folds or flutes. When the sun rotates, an observer near the ecliptic will alternately lie above and below the current sheet and will see a changing solar sector pattern. The inset above shows a meridional cross section with the earth below the current sheet (Figure courtesy of S. I. Akasofu).

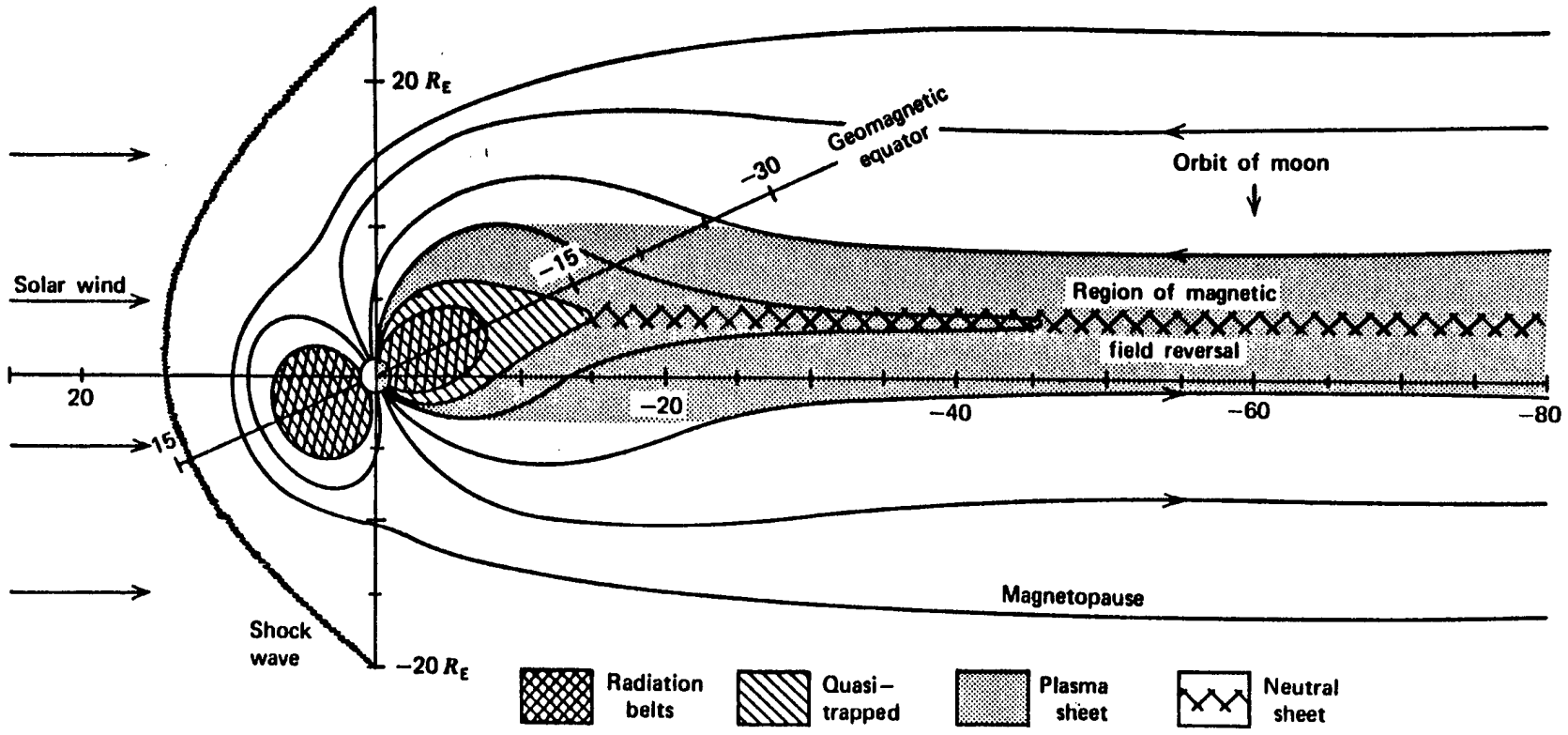


Figure 23. The magnetosphere of the earth, meridian plane view. The time is summer in the northern hemisphere, (Haymes[6], p.186).

gamma is a commonly used unit of magnetic field strength equal to one nanotesla). Sq appears in the northern hemisphere as a decrease in the horizontal component near local noon (see Figure 24). Currents induced in the ionosphere due to the tidal influence of the sun on the atmosphere and ionosphere are the cause of this current system which decreases (increases) the main field at sea level in the northern (southern) hemisphere. This diurnal variation changes the magnetic field at sea level by about 0.05 percent. The system consisting of sun, interplanetary space and magnetosphere is extremely dynamic. More than 97 percent of the time there are variations in the magnetic field over and above the solar-quiet Sq variation.

The discovery that the earth is a large magnet was made by William Gilbert (in Chapman and Bartels[5]) who published (in Latin) "De Magnete" in 1600. In 1635 Gellibrand (in Chapman and Bartels[5]) found that this magnet is not constant but changes slowly with time. Edmund Halley (in Chapman and Bartels[5]) made the first extensive ocean magnetic survey and produced the first world magnetic chart inn 1835. Gauss (in Chapman and Bartels[5]) made the first spherical harmonic analysis of the earth's field and determined the major coefficients. He also set up the first magnetic observatory at Gottingen, Germany.

Thus, systematic observations of the terrestrial magnetic field have been made for centuries. It was recognized quite early that there is great variability in the bumpiness or degree of complication in the magnetogram, the graphic presentation of the variations of the magnetic field, at a given station. An index of magnetic activity 'C' (international daily character with range 0-2) on a daily time scale was used for many years, but since 1932 this "magnetic character figure" has dwindled in importance. Several indices with a finer time scale than one day are now in existence. The most widely used such index is K_p, which is calculated every three hours of Greenwich time; thus there are 8 distinct values of K_p for each day. The subscript "p" was intended to stand for planetary, but the index in fact responds primarily to auroral activity. The index is defined for each 3-hour interval on the basis of the largest value, R, of the 3-hour ranges in X, Y, D, or H, where the range is the difference between the highest and lowest deviation from the regular "quiet" daily variation, The K value for a given value of R is found from a table in which the location of the station has been taken into account to permit comparison between K

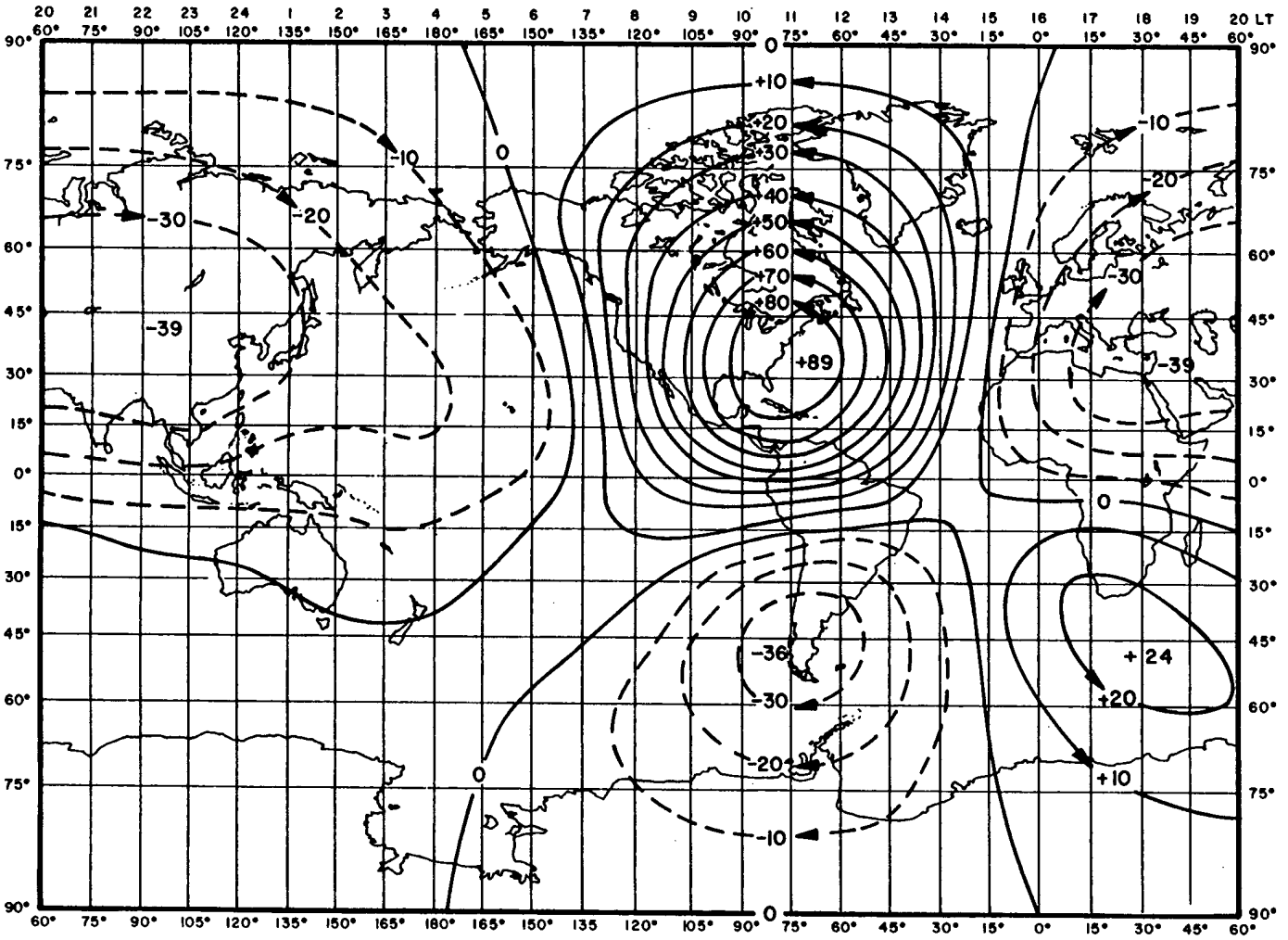


Figure 24. The global ionospheric systems responsible for the Sq effect. A current of 10,000 Amperes flows between ovals, (Haymes[6], p.240).

values from different stations. The K_p index is based upon K values from 12 stations between geomagnetic latitude 48° and 63°, selected for good longitude coverage. The average K value so obtained is quoted to thirds of a unit by appending a "-", "0", or "+" to the integer between 0 and 9 . Thus there are 28 possible values of K_p : 0_o, 0_+, 1_-, 1_0, 1_+, 2_-, 2_0, 2_+, ... 8_+, 9_-, 9_0.

Air Weather Service (AWS) has its own network of 8 magnetometer stations (Prochaska[7]) in the northern hemisphere covering the longitude range of the CONUS sector (Alaska, US to UK) for real time determination of the AWS K index (similar to that of K_p).

Records of K_p are complete beginning in 1932. The K_p index is roughly logarithmic in the range R. Another widely used index, derived from the K_p indices is the ap index which is roughly linear in the same range; the relation between K_p and ap is shown in Figure 25. Ap is the daily average of ap. Since we have 8 values of K_p for every day from 1932 through 1990 inclusive, we have a total of 172,400 values with which to do " statistics". The simplest calculation requires counting the number of times each possible value of K_p occurs. Figure 26 shows the relative frequency, the count of each K_p divided by the total number, 172,400 plotted versus K_p. We see that the quietest value, 0_o, occurs slightly more than 2 percent of the time, while the stormiest, $K_p = 9_o$, has occurred only 25 times in 59 years. The median value is 2_o; there are as many values less than 2_o as there are greater than 2_o. The mean value, 2.6537, is very close to a K_p of 3_-.

Cumulative frequencies are given in Figure 27 showing what fraction of the time a given value of K_p is exceeded, or how many times a value less than or equal to a given value is attained. Figure 28 shows the relative frequency of ap, with a median of 7 and mean of 14.7. The inset expands the scale and shows the results up to ap=50.

Figure 29 shows two variables plotted versus year from 1932 to 1990 inclusive. The curve marked the 'SUNSPOT NUMBER' presents the unsmoothed annual mean of sunspot number, R_z. The other curve (with symbols) represents the yearly count of active days, defined by the Air Weather Service as having a daily K_p sum greater than or equal to 25. On this 'MEAN K_p' curve the occurrence time of sunspot maxima and minima are marked to show the differences in the variation of K_p and sunspots with time. Figure 30 shows, in the same format, the sunspots

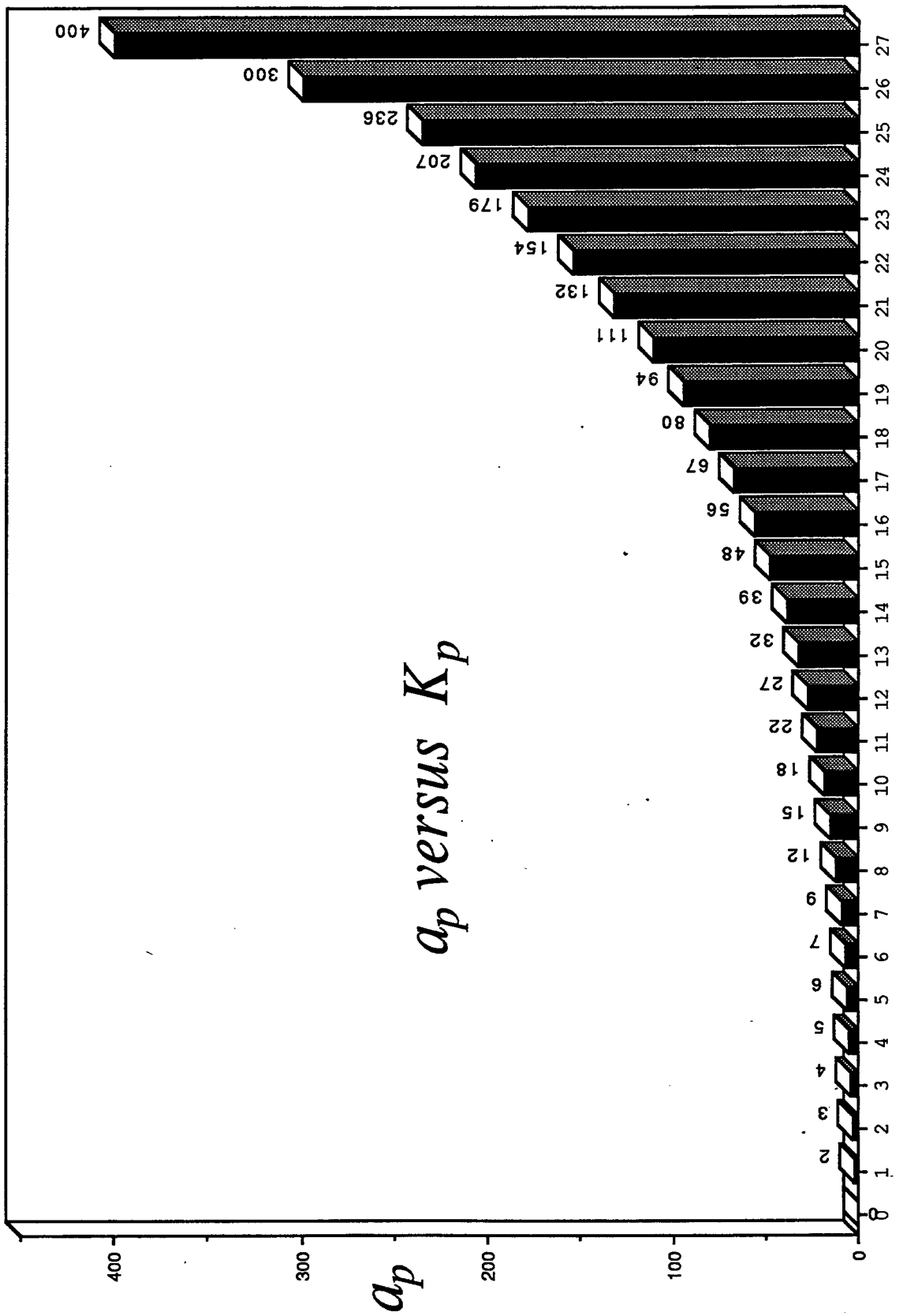


Figure 25. Relation between ap and K_p.

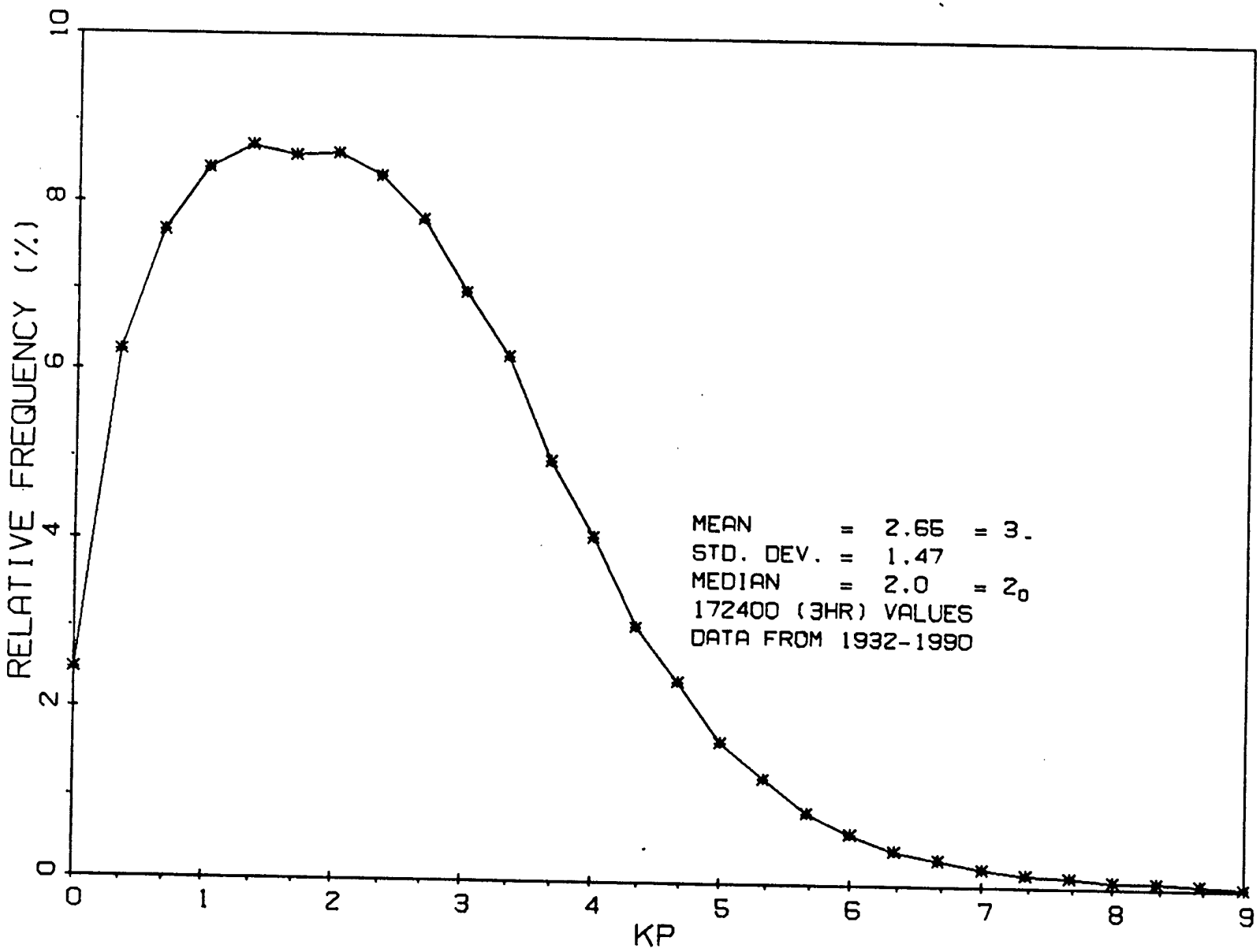


Figure 26. Relative frequency of occurrence of K_p.

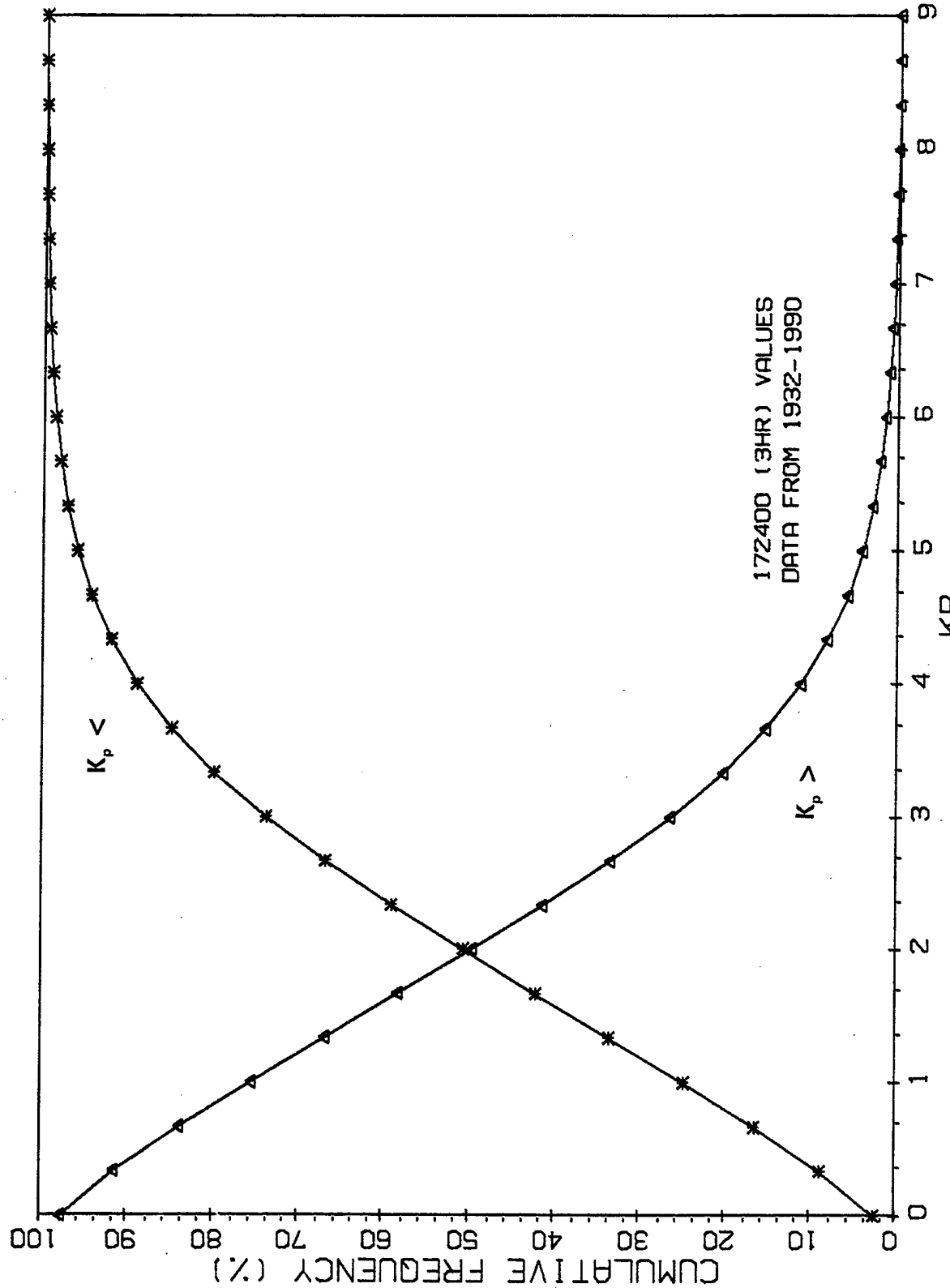


Figure 27. Relative cumulative frequency of occurrence of K_p.

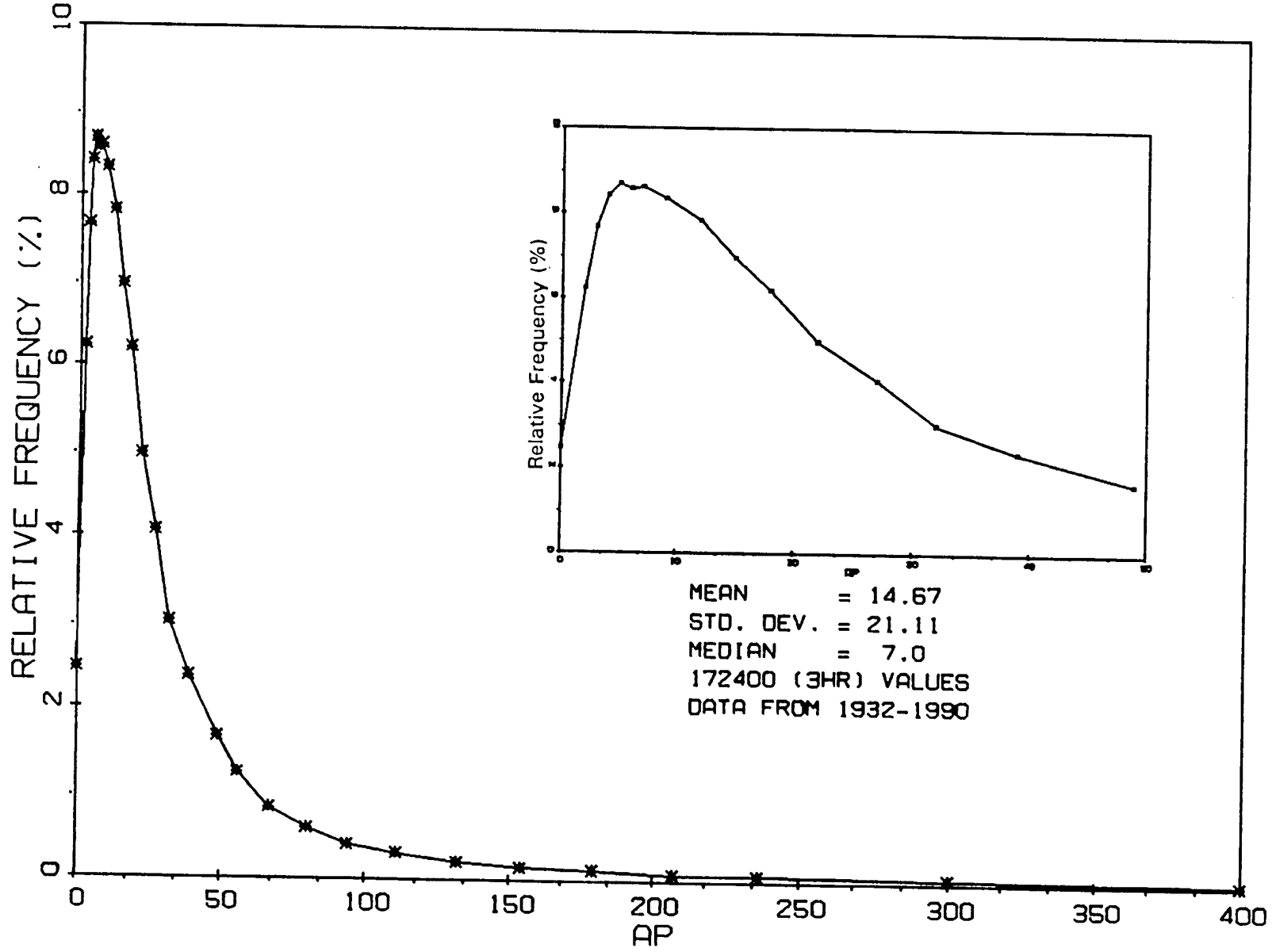


Figure 28. Relative frequency of occurrence of ap.

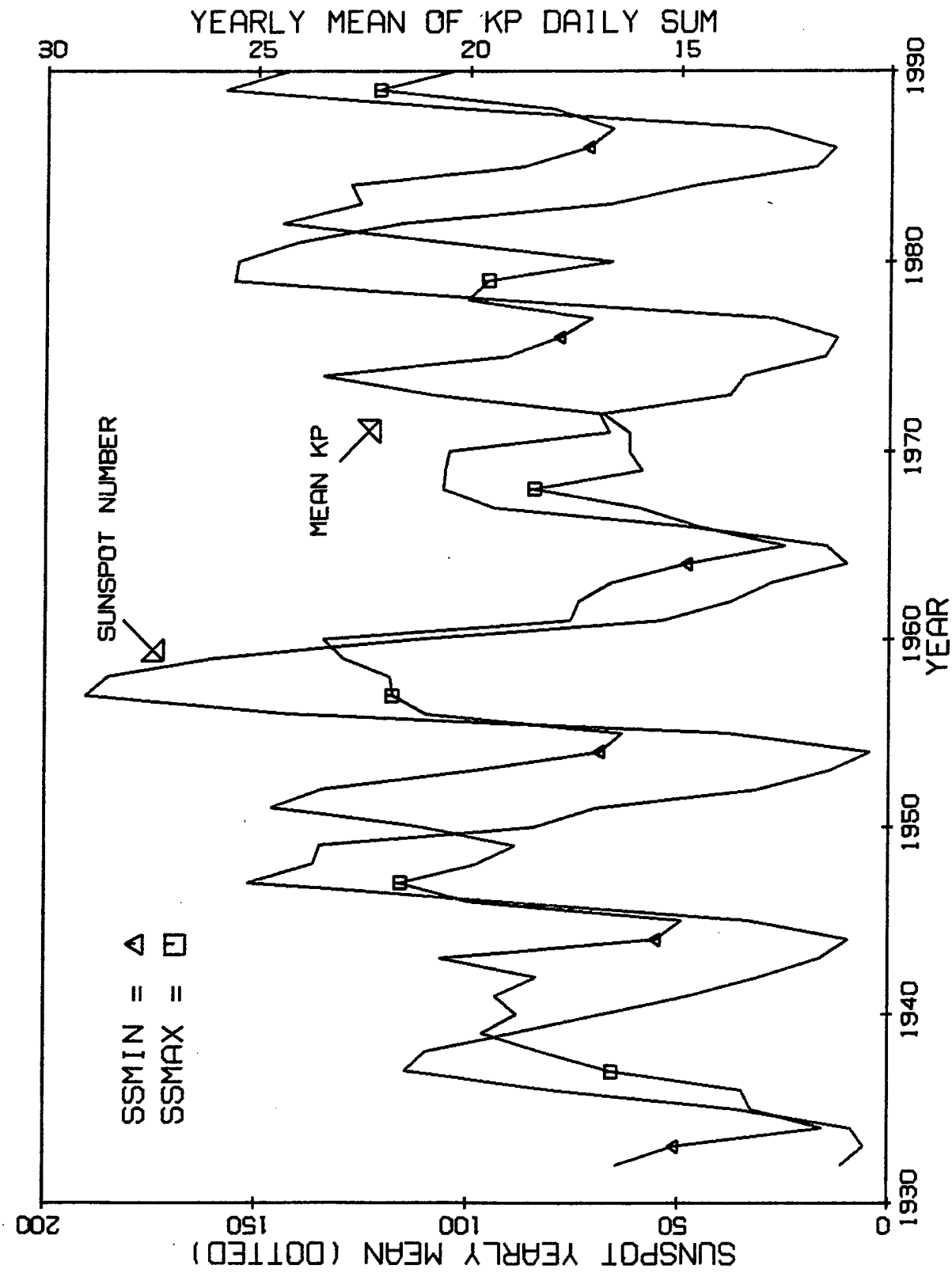


Figure 29. Comparison of mean annual sunspot with mean daily sum of K_p.

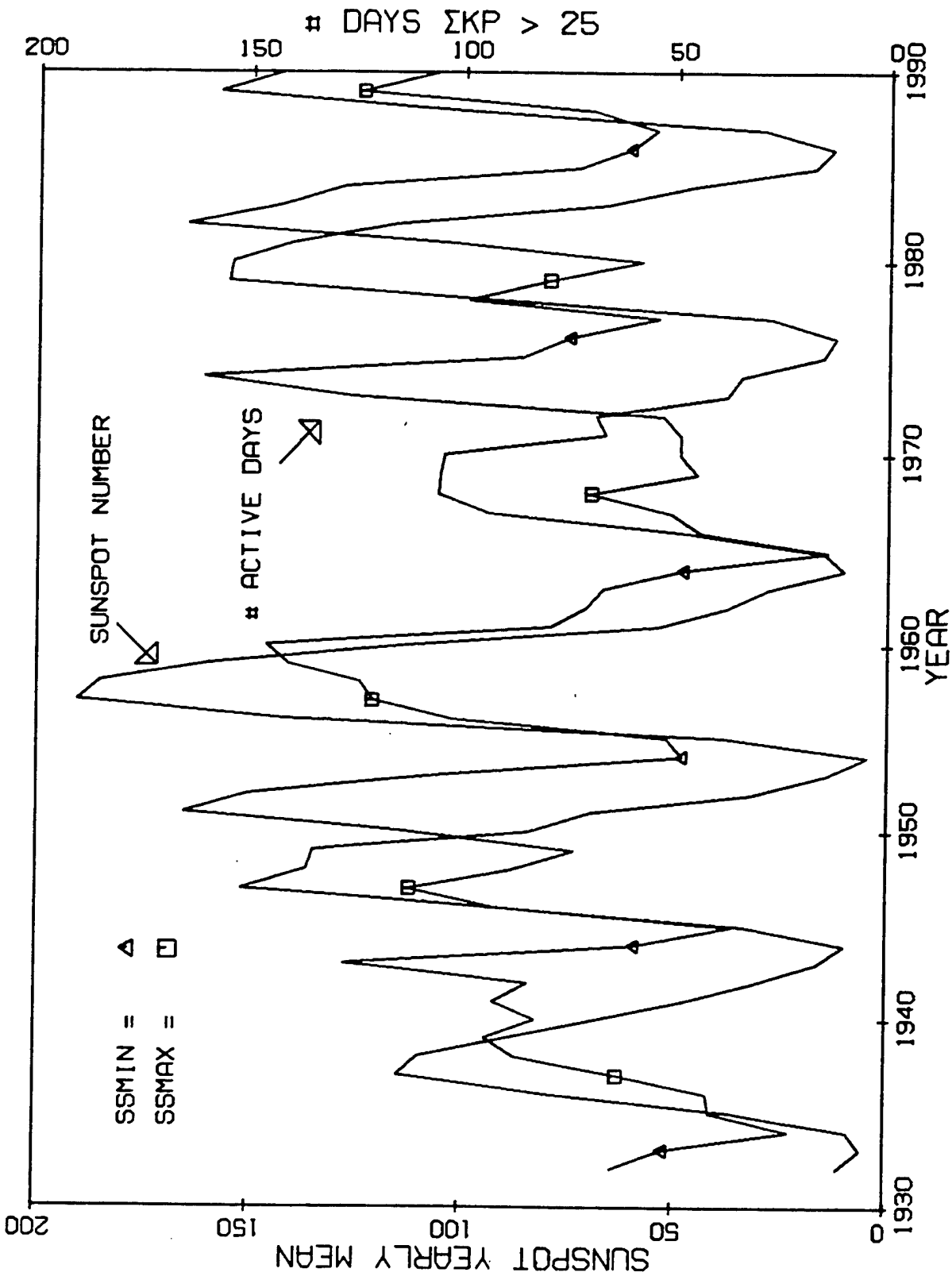


Figure 30. Comparison of mean annual sunspot with days with $\Sigma K_p > 25$.

and the yearly mean value of K_p. Comparing these two figures by eye reveals a startling similarity. It turns out that the number of active days is very nearly proportional to the annual mean of K_p. The figures show that minima in sunspot activity are nearly always accompanied by minima in magnetic activity but the converse is not true. In fact, during some years, notably 1946 and 1948, minima in magnetic activity correspond to the increasing phase of solar sunspot activity. Also some minima of magnetic activity actually correspond to maxima of sun spot number (see for example, 1980). Thus there is no simple relation between solar (sunspot) and geomagnetic (K_p) activity.

3. DISTURBED IONOSPHERE

3.1 Short to Medium Duration Disturbances (Hours to Days)

There is sufficient variability in the "normal" ionosphere with time of day, season, altitude, latitude, and sunspot number- that it is almost unfair to throw additional variables into the "ionospheric pot". The sun is responsible for the ionosphere's existence, and is ultimately responsible for virtually all the ionospheric disturbances that can interrupt OTH operations. We will now examine each of these ionospheric disturbances.

3.1.1 SUDDEN IONOSPHERIC DISTURBANCE (SID)

The SID becomes apparent a few minutes after some large solar flares and lasts as long as the active phase of the flare (a few minutes to an hour). It is the result of a strong increase in X ray flux associated with the flares. The ionospheric effect is a sharp fading of long distance HF radio communications (and in some instances OTH radar returns) on the sunlit side of the earth. A strong enhancement in electron density often occurs in the D and lower E layers, which increases absorption of HF rays that pass through these layers. As a result, an HF user must employ a higher frequency in order to propagate a signal through the ionosphere (higher LUF, the lowest useable frequency, see report by G. Sales PL-TR-92-2123) or more typically, (especially for fixed point to point communication and for the OTH fixed barrier location operating scenario,) the user must wait for the SID to pass.

3.1.2 POLAR CAP ABSORPTION EVENTS (PCA)

A PCA event is a widespread, long-lived increase in absorption confined to the polar ionosphere. Like SIDs, PCAs are caused by solar flares, but there is a completely different mechanism. Whereas the SID is caused by enhanced solar radiation, the PCA is brought about by high energy protons (greater than about 5 MeV) generated by the solar flare, traveling to the earth in 20 minutes to a few hours. These particles are guided by the earth's magnetic field lines into the polar caps. Here, they give up their energy by ionizing neutral particles at the altitude of the D layer. The resulting increase in electron density strengthens and lowers the polar D region, at times down to 40 km, and increases the absorption of HF signals passing through this region. As with an SID, the result is an increase in the LUF on polar sky wave paths. With major solar flares, the LUF may even exceed the MUF (maximum useable frequency) of higher ionospheric layers (E, F_1, and F_2) making polar HF transmissions impossible. Polar cap absorption events are generally of little importance to the current midlatitude installed OTH radars, except for the fact, that they are often accompanied by magnetic storms.

3.1.3 GEOMAGNETIC/ IONOSPHERIC STORMS

Geomagnetic disturbances have a variety of possible causes: flares, disappearing solar filaments, passages of solar sector boundaries in the interplanetary magnetic field, solar coronal holes, etc. After the magnetosphere is disturbed, it is difficult to predict the timing or magnitude of the ionospheric disturbance. Considerable energy is injected into the auroral oval during a geomagnetic disturbance. Some of the energy is due to precipitating particles, while the remainder is due to Joule heating by storm-associated currents. This energy input sets up large scale neutral atmosphere convection, disrupting the normal chemical composition of the atmosphere. Overall, when an ionospheric storm begins, the total electron density on a global scale will decrease to some minimum value and then slowly recover. Successive storms may cause additional reductions in electron density, but some absolute minimum value will exist at any point. During the recovery phase, which may last up to a day at solar maximum and up to several days at solar minimum, daytime increases due to solar radiation will initially overbalance nighttime losses as the ionosphere refills. This reduced electron density period lowers MUFs and

reduces barrier depths for the OTH radar, while the recovery phase increases MUFs, expands the barrier depths and enhances the radar detection performance. The ionospheric average over this cycle produces a net increase in total ion/electron content compared to undisturbed periods. The overall effect on the radar is its reduced performance and is seen as lowered clutter to noise ratio on aircraft/ targets, rendering them undetectable.

One class of geomagnetic disturbances is the auroral substorm. Auroral substorms are due to particle precipitation into the high latitude ionosphere. They cause strong radio wave absorption on auroral paths, intense sporadic E (see below), and strong irregularities. Auroral substorms occur relatively frequently, especially during periods of high solar activity; several may occur in a 24 hour period, each lasting for one to two hours.

3.1.4 TRAVELING IONOSPHERIC DISTURBANCES (TID)

TIDs are short-term disturbances produced by wave-like perturbations related to atmospheric gravity waves. Normally found in the F layer, these waves usually originate at high latitudes and spread equatorward. TIDs affect OTH radar operations by producing a wavelike lowering and raising of the reflection height of the ionosphere, resulting in a change in location of the barrier. TIDs can produce Doppler variations in radar returns, resulting in velocity transients in aircraft tracks.

3.1.5 SMALL SCALE IRREGULARITIES

Ionospheric clutter, as detected by OTH, results from direct backscatter of radar energy by intense small scale ionospheric irregularities. These irregularities are most common in the auroral and equatorial regions but occur also during equinox and winter nights and during periods of enhanced magnetic activity in midlatitude areas. These irregularities are generally aligned with the magnetic field. The clutter returns are most intense when radar energy encounters such an irregularity while the signal propagation is nearly perpendicular to the long axis of the irregularity, that is, the signal propagation direction is perpendicular to the earth's magnetic field lines in the area of irregularity occurrence. These clutter conditions can be present over a wide area and affect OTH frequently at day/ night transitions, and at night during equinox and winter periods.

3.2 The High Latitude Ionosphere

The high latitude ionosphere is a product not only of the solar radiation and the terrestrial effects described earlier, but also of solar-magnetospheric effects. These result from the interaction of the magnetic field of the earth with the solar wind, the plasma (chiefly ionized hydrogen and electrons) boiled off from the sun. There are two general phenomena of magnetospheric origin, the aurora and ionospheric-magnetospheric convection. Together they produce a multitude of varying and overlapping effects that are permanent features of the high latitude ionosphere. These large effects often produce rapidly varying levels of electron density, irregularities and rapid movement of ionospheric constituents at all levels of the ionosphere and can dominate the ionosphere, particularly at night, in such a way as to make it unsuitable for OTH operation, by limiting HF propagation, or by masking targets with clutter, or both.

OTH radar uses ionized layers in the upper atmosphere to refract radio energy in the HF band back to the ground to detect and track targets (aircraft) at much greater distances than what line of sight allows. Figure 31 shows schematically the ray path between an OTH radar and an aircraft via an ionized layer near 300 km altitude, the F layer. (This format will be used in the following to illustrate some of the effects of the high latitude ionosphere on OTH radar). Because the ionosphere is an integral part of the radar system, a disturbance of the ionosphere results in impairment or enhancement of the system. This section discusses the naturally occurring disturbances in the high latitude ionosphere, and some of the limitations these disturbances can place on OTH radar operation.

3.2.1 THE AURORAL IONOSPHERE

3.2.1.1 Aurora

The aurora borealis (or 'Northern Lights') is the name given to the light produced by particles (electrons and protons) that precipitate from space and collide with atoms and molecules in the earth's upper atmosphere. The process is analogous to a cathode ray tube in a television set in which beams of electrons are accelerated to tens of kilovolts by an electron gun, pass through the vacuum tube and, guided by magnetic fields, strike the screen to produce light. Similarly, auroral particles accelerated by plasma processes follow the earth's magnetic field

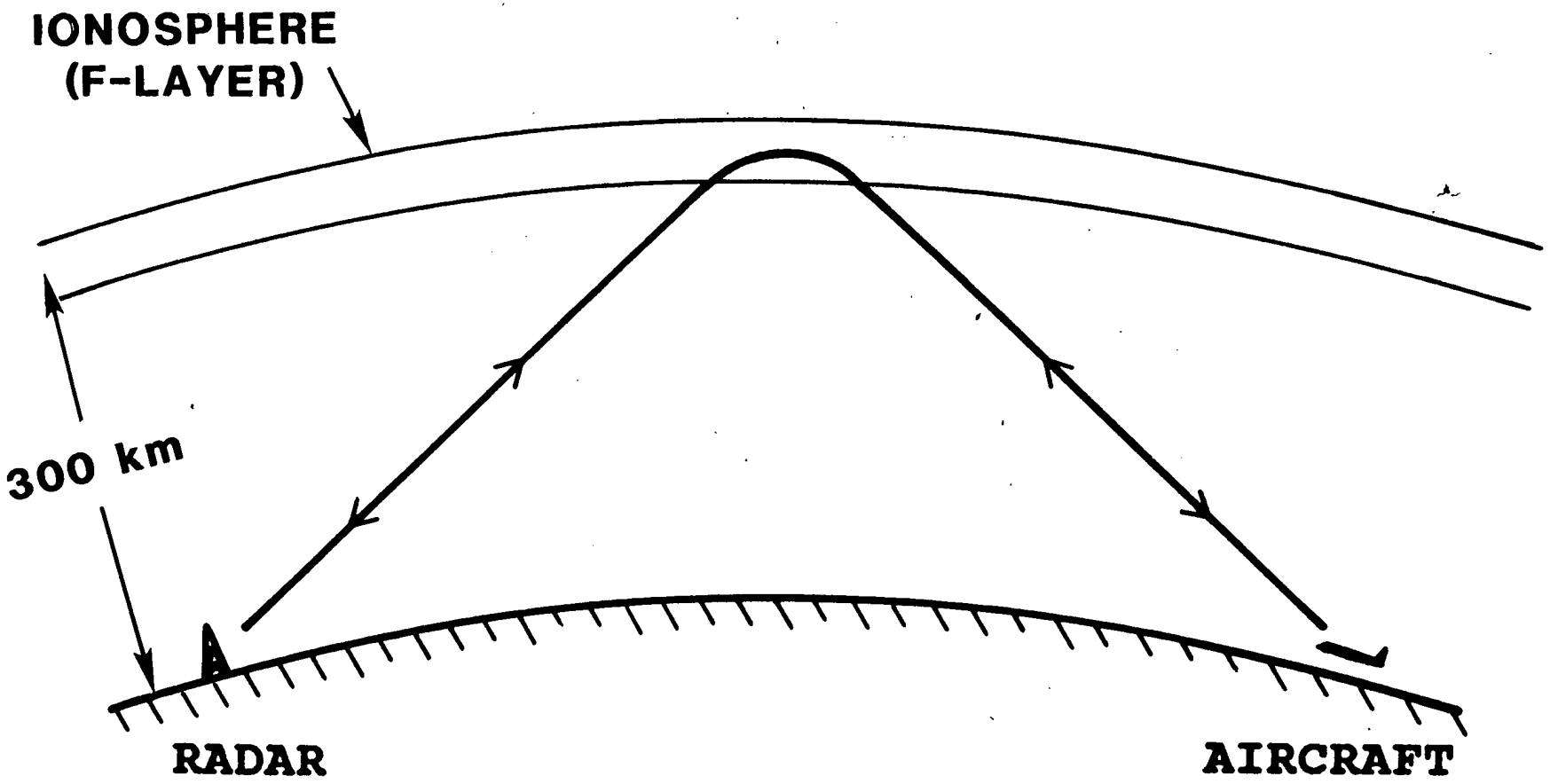


Figure 31. Schematic view of the way OTH uses HF propagation to extend its range beyond line of sight.

lines through the low density space environment until they strike the higher density of the upper atmosphere. Unlike the television, auroral emissions consist of atomic line and molecular band emissions such as those occurring in gas discharges within neon signs and fluorescent lights.

Viewing the aurora from below on the 'screen' of the upper atmosphere, one sees a very chaotic and confusing picture since only a small part of the entire pattern is visible at one location. However, seen as a whole from an altitude of 20,000 km by the Dynamics Explorer I satellite, the aurora has a very regular pattern: an enormous circular band some 4000 km in diameter called the auroral oval, shown in Figure 32. In the figure the midnight sector of the oval is the wider part at the lower right, and the midday sector is the narrow part at the upper left. The broad region at the left of the image is the sunlit daytime atmosphere. Ultraviolet emissions are imaged here, but their intensity and location correspond closely to visible emissions. The model of the oval will be described in detail below.

In the process of slowing down and stopping, these auroral particles not only produce light but also ionize the atmosphere. The altitude at which the maximum ionization is produced depends on the initial kinetic energy of the incident particles: the more energetic the particles, the deeper they penetrate into the atmosphere, hence the lower is the altitude of the resultant ionization.

Figure 33 illustrates the situation for the major classes of auroral electrons. From right to left, electrons of about 0.1 keV (1keV = one thousand electron volts) produce maximum ionization at heights of 250 to 300 km (which is the altitude region designated as the F region); electrons of 1 to 10 keV ionize principally between 100 and 150 km altitude (the E region); those with energies greater than 40 keV ionize below 90 km (the D region). The 1-10 keV electrons produce the most visible aurora and account for the auroral emissions recorded in Figure 32. All three classes produce ionization that affects HF communication in different ways, most of the time, adversely.

3.2.1.2 E Region - The Auroral Oval

The most visible optical emissions are those resulting from 1 to 10 keV electron precipitation that ionizes the E region. These emissions, which form the auroral oval shown in

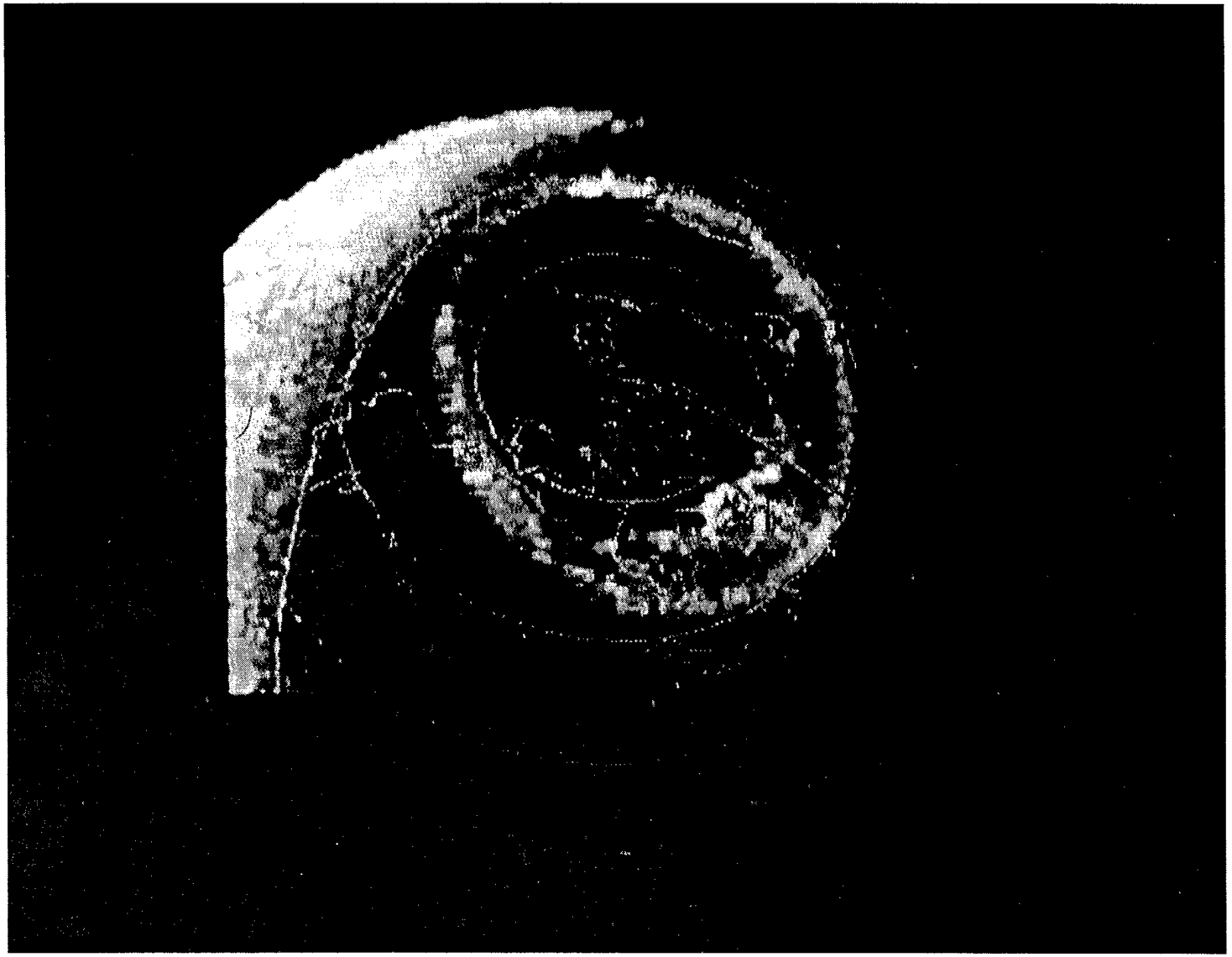


Figure 32. The auroral oval over the northern hemisphere. The oval is the circular band, the wide part of which (at lower right) is midnight, the narrow part (at upper left) is noon. The broad area of light at the left is the sunlit atmosphere. Continental outlines and parallels of magnetic latitudes at 15° intervals are superimposed on the image of ultraviolet emissions recorded by the Dynamics Explorer 1 satellite from an altitude of 20,000 km at 0241 UT on 8 Nov 1981, (Courtesy of L. A. Frank).

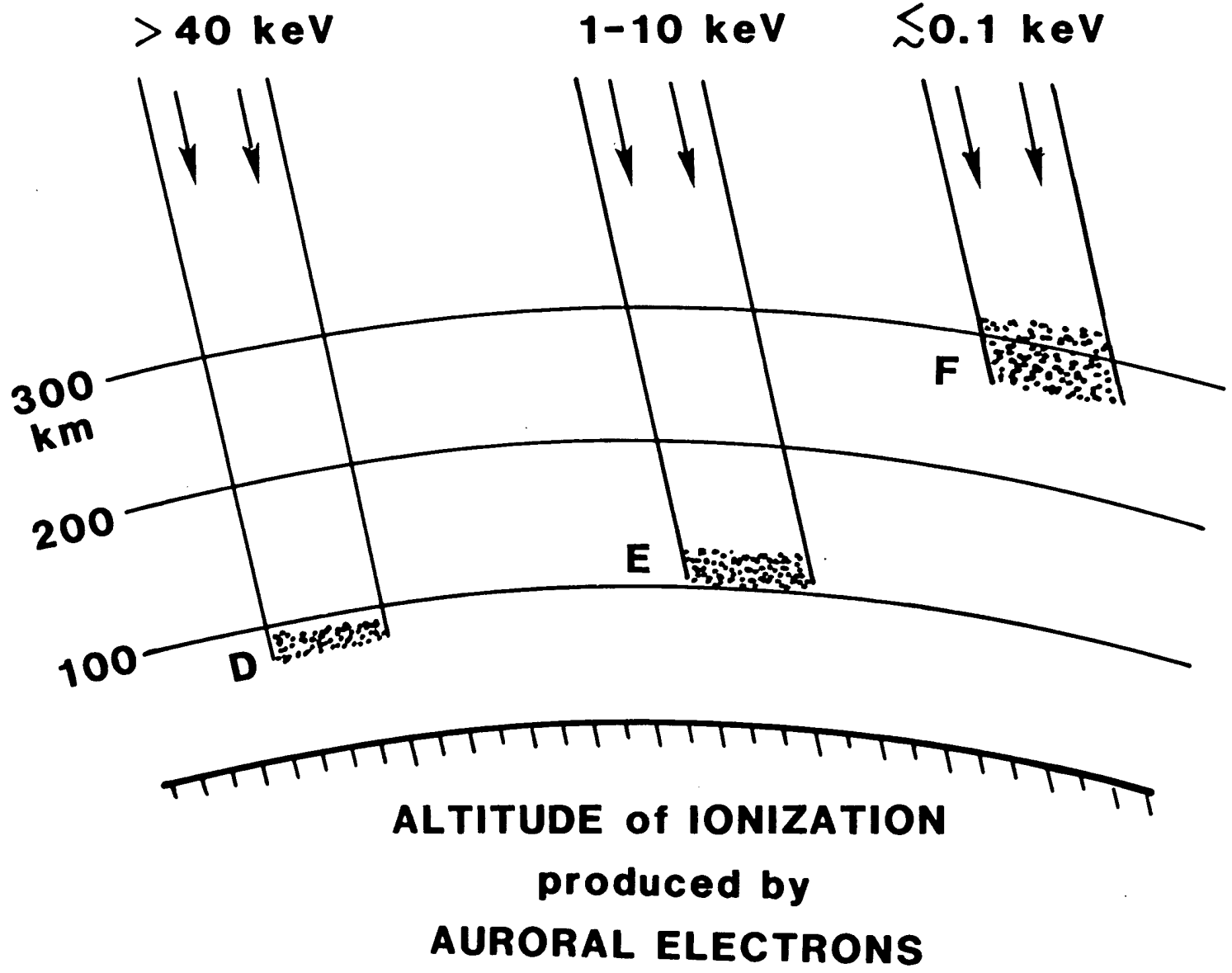


Figure 33. The relation between the kinetic energy of precipitating auroral electrons and the approximate altitudes at which they produce maximum ionization.

Figure 32, are collocated with the ionization in the E region. However, most auroral ionospheric phenomena and their effects are not confined to the oval. Even so, the oval can serve as an illustration of some general properties that are common to all the phenomena, as well as a convenient frame of reference for those that occur outside the auroral oval boundaries.

A model of the oval by Feldstein and Starkov[8] projected on the earth as plotted in geomagnetic latitude/ geomagnetic local time (GMLAT/ GMLT) is roughly circular in shape but eccentric about the magnetic pole so that it is at lowest latitude near midnight and highest latitude near noon (Figure 34).

The coordinates in which the oval resides are stationary with respect to the sun, and the earth rotates beneath it. The resulting diurnal variation viewed from the earth during the course of 24 hours is illustrated in Figure 35 where the map of the earth is plotted in the coordinates of geomagnetic latitude/ geomagnetic longitude (GMLAT/ GMLONG). Also shown as the triangular regions are the northernmost coverage sectors of the East Coast and West Coast radars, (note that in Figures 35, 36 and 37 the range of ECRS coverage is shown as 1800 nmi; since March 93, in all sectors this range has been increased to 3000 nmi and all the sectors have been rotated southward by 15°). The oval intrudes into a large part of the east coast sector during the nighttime hours (00 to 06 UT), whereas the west coast sector is relatively unaffected.

The oval varies in size as a function of magnetic activity, contracting poleward and narrowing in width during quiet conditions, expanding equatorward and widening as the magnetic activity increases (Figure 36). Quiet, moderate and active conditions are represented by magnetic index Q, an index determined from magnetometer readings at an auroral latitude of 65° in the midnight meridian at 15 minute intervals. Currently 'Q' is treated as a morphological index and is inferred from the 'Q_e' index derived from electron precipitation measurements by DMSP satellites.

The combination of probability of oval size and of the duration of the oval within the coverage area is shown in Figure 37. In the black area in the northeast corner the oval is present 100 percent of the time. The contours at 75, 50, and 25 percent show the percentage of time during a 24 hour period that the aurora is present poleward of these boundaries. In the shaded area in the southeast corner, the oval intrusion is statistically 0 percent.

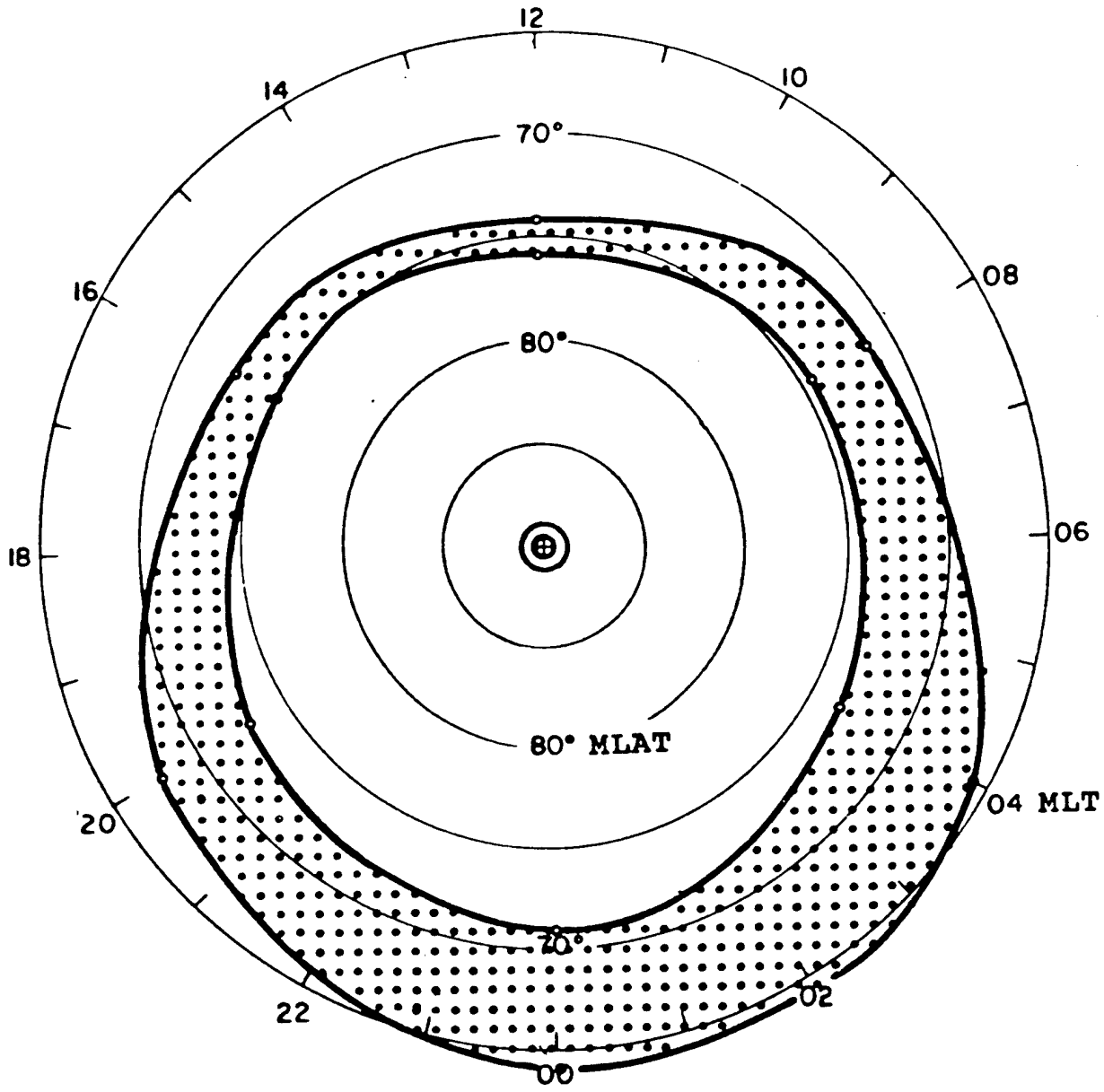


Figure 34. The auroral oval model of Feldstein is shown at a typical size (magnetic index Q=3). Coordinates are magnetic latitude/ magnetic local time (GMLAT/ GMLT) in the Corrected Geomagnetic System. Note the similarity to the image in Figure 32. The oval is the locus of discrete aurora, sporadic E and irregularities that cause clutter in OTH radar.

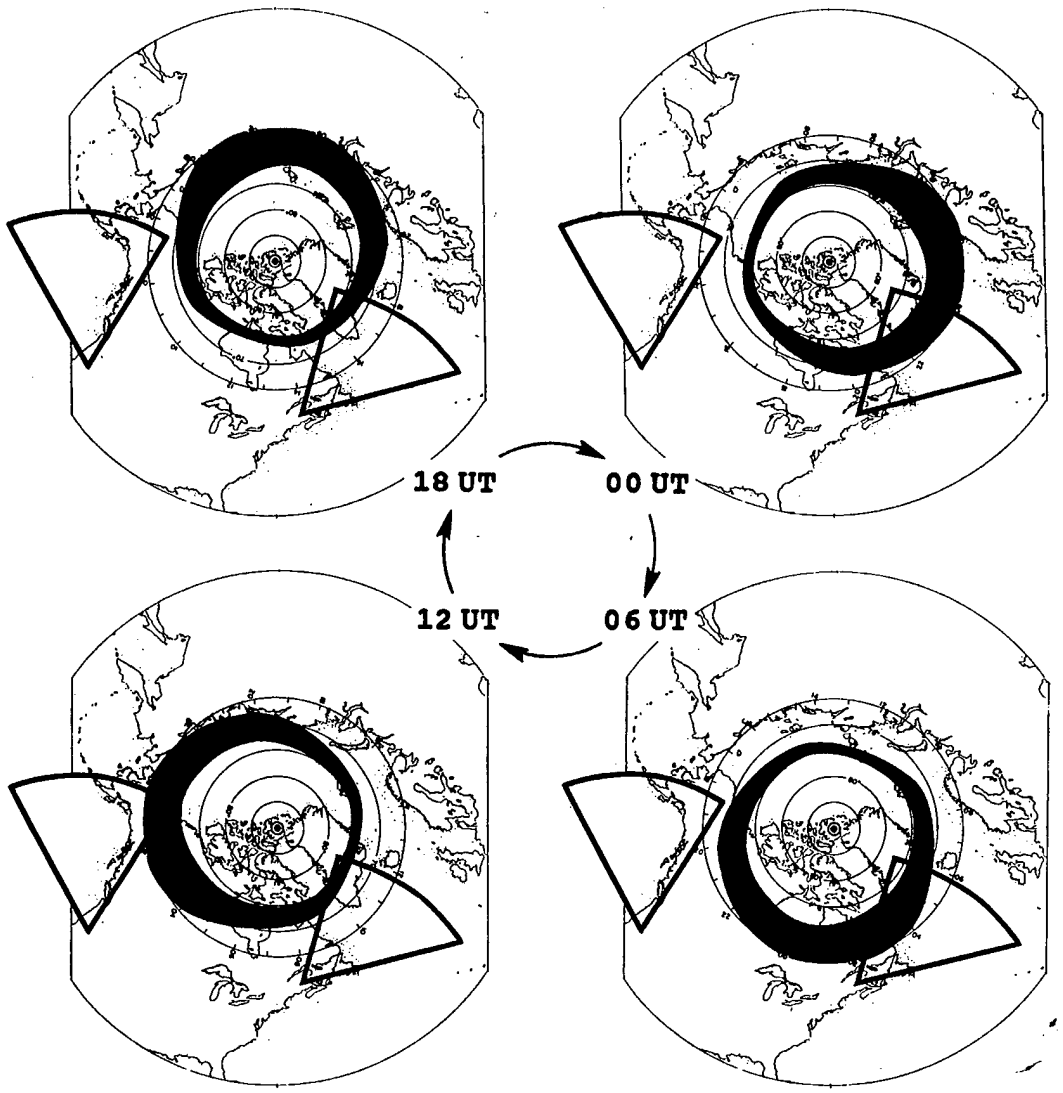


Figure 35. Diurnal variation of the auroral oval in relation to OTH coverage areas. The auroral oval superimposed on the northern hemisphere is mapped in magnetic latitude/magnetic longitude (GMLAT/GMLONG). The northernmost OTH coverage areas are shown as the as the two pie-shaped sectors. Four Universal Times (UT's) are shown.

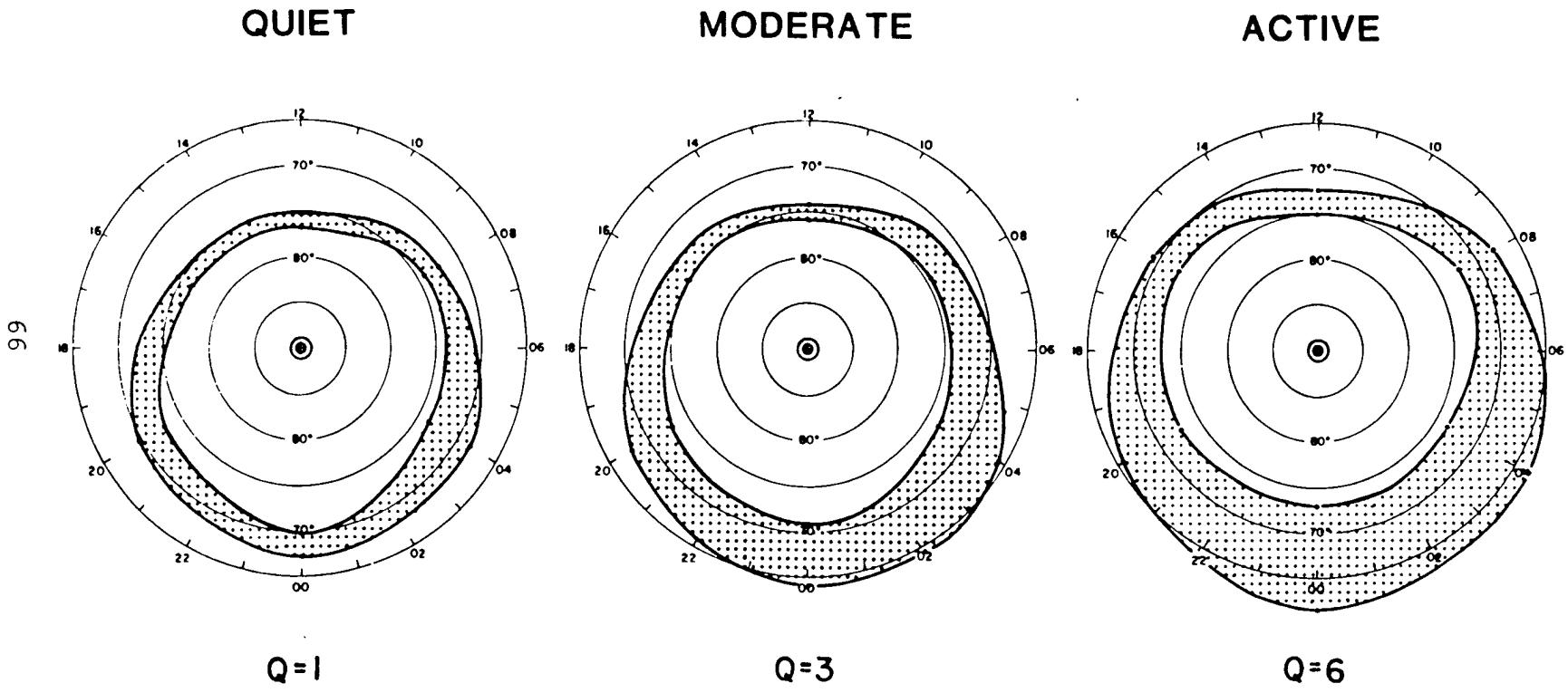


Figure 36. Variation of the auroral oval with activity. The auroral oval is shown under three conditions, quiet, moderate and active, corresponding to the three values of the magnetic index, Q, of 1, 3 and 6 respectively.

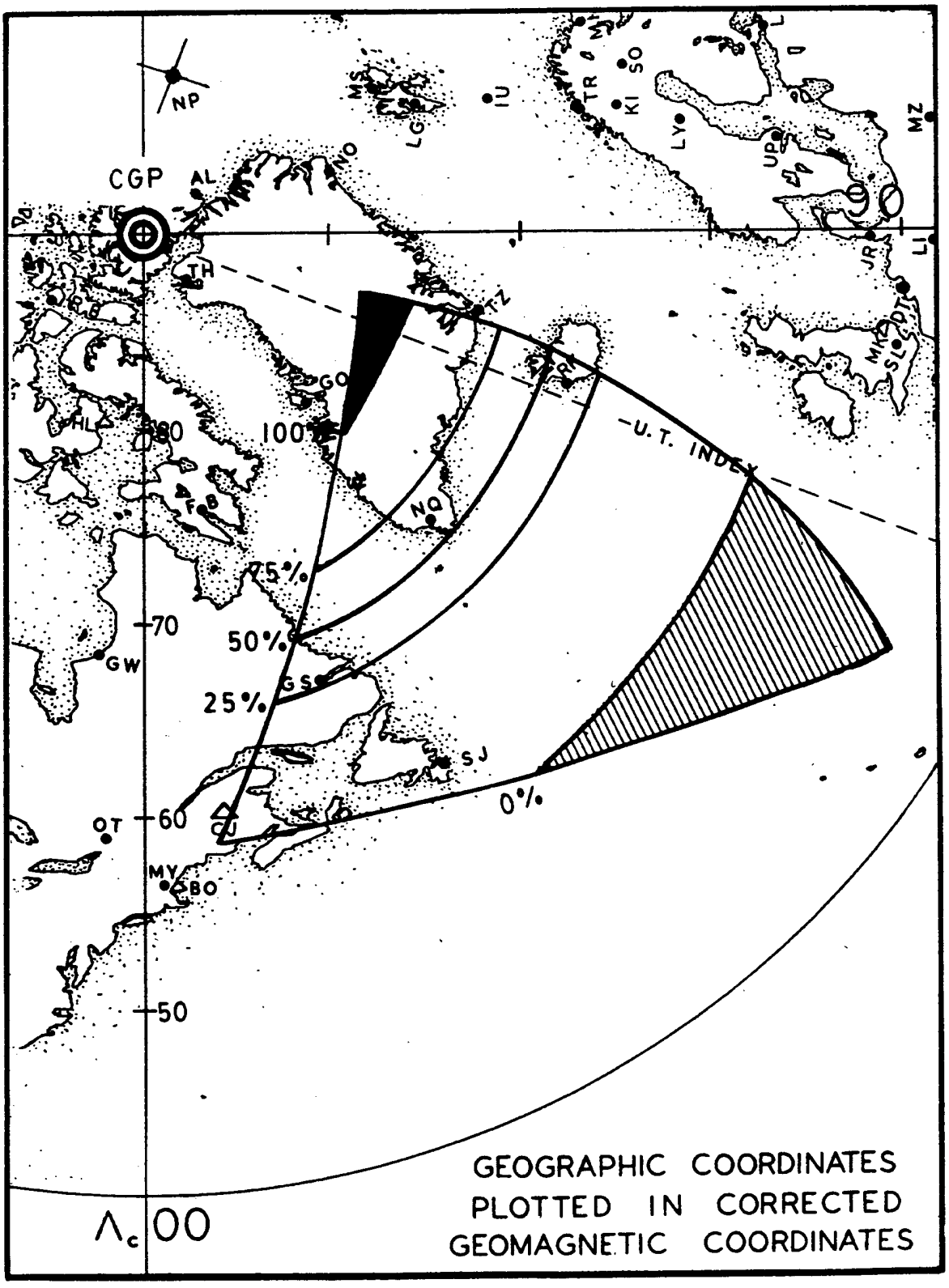


Figure 37. Percentage of time that the auroral oval is present within the coverage area, (Dandekar[9], p.23)

Section 3.2.2.2 will show that other phenomena (auroral absorption and the convective effects of trough and clutter) can extend equatorward of the oval. Accordingly the contours for these effects can be expected to extend somewhat farther (southward) into the coverage area than the oval itself.

3.2.1.3 E Region Ionization Effects on Propagation

There are two general types of E region auroras, discrete and continuous (or diffuse). The discrete aurora consists of a large variety of forms such as visible arcs, bands, and rays that are commonly described as 'the aurora', the locus of which is the auroral oval itself. The resulting ionization, which is spatially limited and often rapidly varying, gives rise to auroral sporadic E (E_{sa}) and irregularities in electron density, the most intense of which are associated with the most active auroral forms. In addition the F region at the location of the oval is generally so disturbed that it cannot support HF propagation.

The continuous or diffuse aurora, though seldom visible to the naked eye, is the most stable and permanent feature of all auroras. It is several hundred km in width, is spatially continuous and slowly varying in time. The locus of this aurora is a band lying in the equatorward part of the auroral oval in nighttime, and equatorward of the oval in the daytime. This is shown together with the oval as the shaded region in Figure 38.

The continuous/diffuse auroral precipitation gives rise to an auroral E layer (Ea) that is similar to that produced by the sun. However currents flowing within the auroral E layer generate small-scale (hundreds of meters) irregularities. The irregularities in the auroral E layer and especially in the discrete aurora strongly scatter HF energy to produce ionospheric spread clutter that can mask the Doppler returns from aircraft. One such case occurs when an aircraft is beneath E region auroral irregularities that scatter so as to be received as clutter at the same range as the aircraft (Figure 39).

In general this spread clutter reduces radar performance in sectors looking in the direction of the auroral oval, which are typically the northern 2 to 4 sectors of Segment 1 from 00 UT to 06 UT as seen in Figure 35. Careful selection of frequencies and waveforms is required to mitigate these auroral effects, and mitigation may not be always successful.

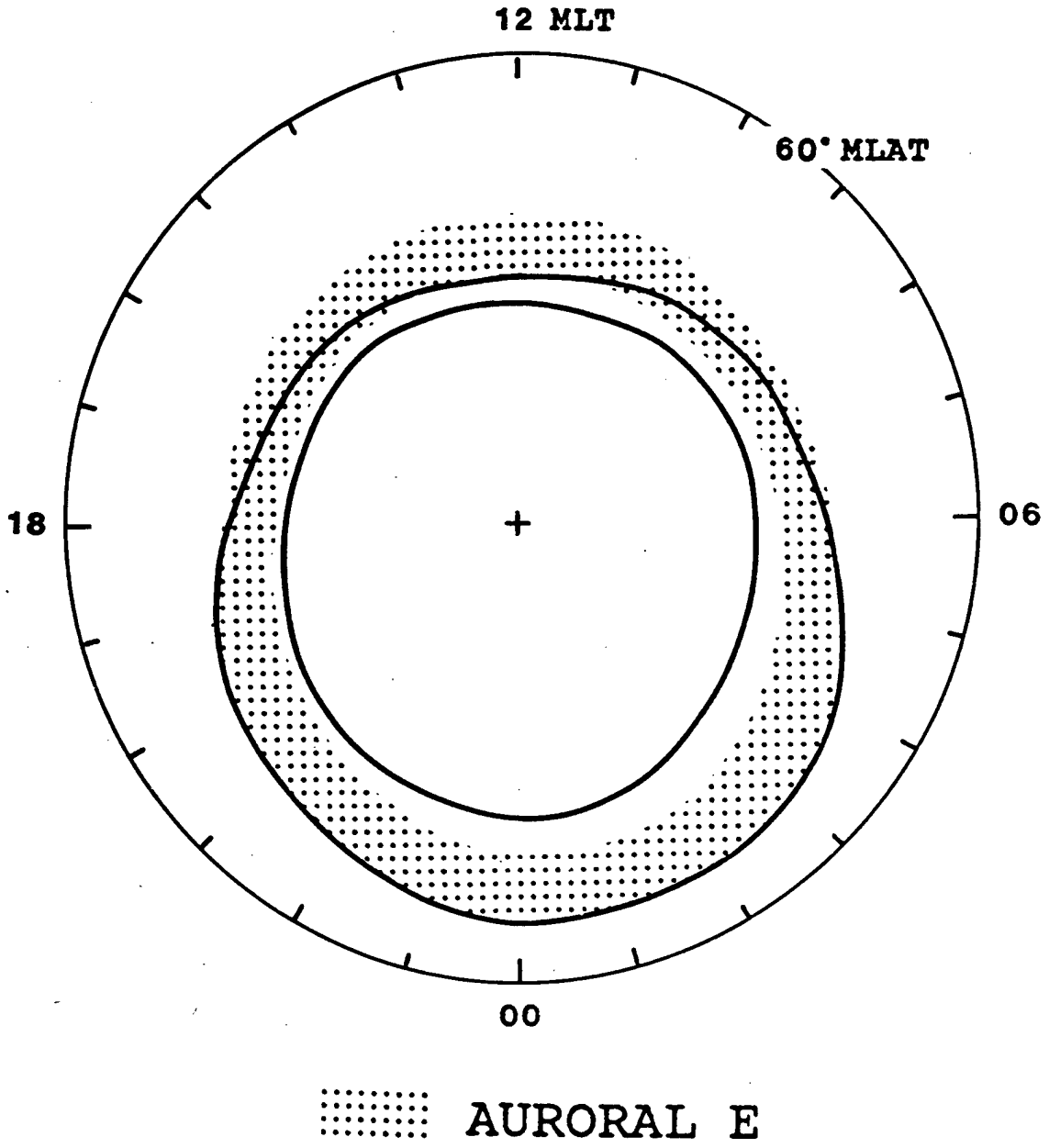


Figure 38. Location of the auroral E layer. The region of precipitation of 1-10 keV electrons that produce the continuous/diffuse aurora is shown schematically in relation to the auroral oval. This auroral precipitation produces the auroral E layer in which are imbedded irregularities that scatter HF energy to cause clutter.

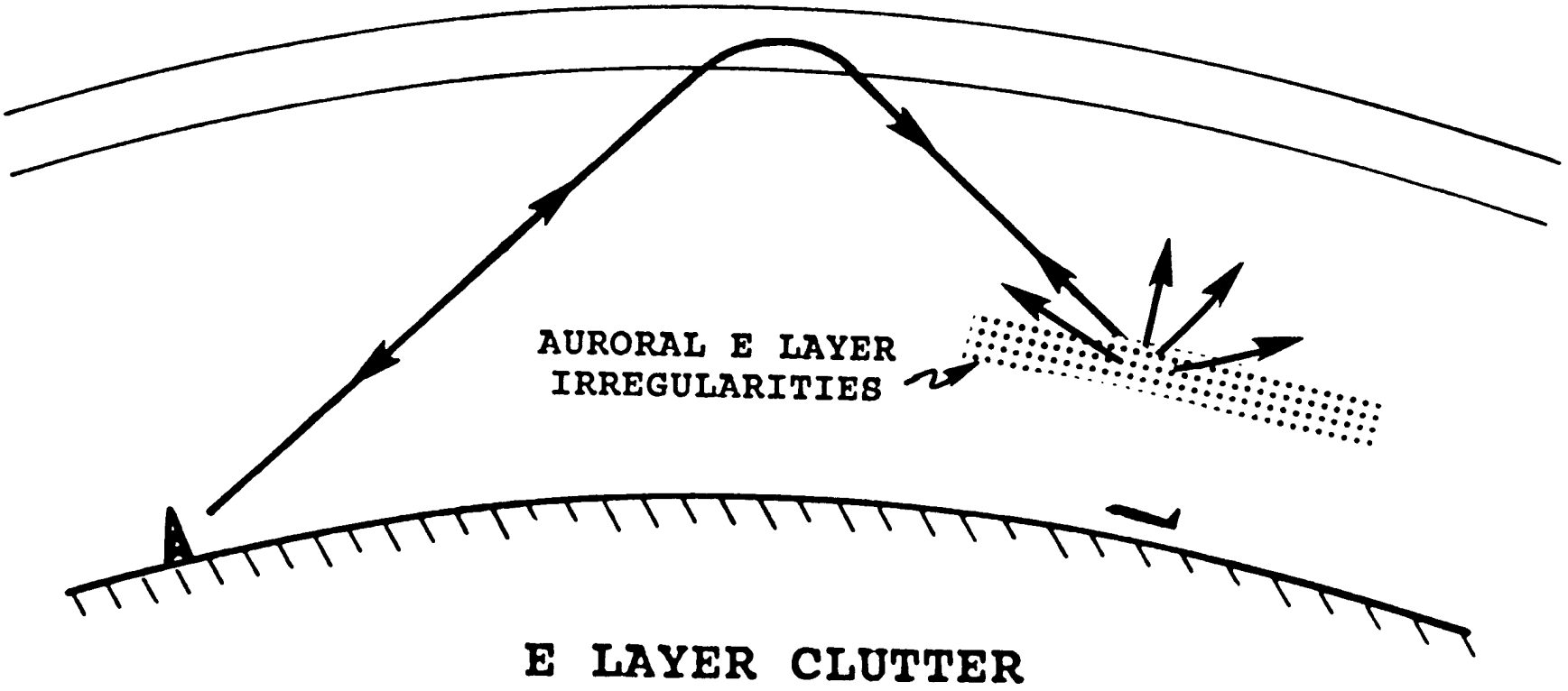


Figure 39. Schematic showing an example of scattering of the OTH signal by irregularities in the auroral E layer that can produce clutter at the range of the aircraft.

In addition the radar, when looking in directions away from the oval, can detect auroral clutter in side lobes of a beam. Even though the efficiency of the side lobes is small, these irregularities are such efficient scatterers and occupy such a large extent in azimuth and range that they can seriously affect the radar performance, particularly during disturbed conditions.

3.2.1.4 D Region Auroral Absorption

A second important category of aurora results from electrons with energy 40 keV and greater that precipitate to produce ionization in the D region (Figure 33). Like the continuous aurora, the light produced from this precipitation is generally too faint, widespread and unstructured to be visible. The resulting ionized D layer is similar to that produced by solar radiation, with similar effects on propagation: radio frequency energy is absorbed with a $1/f^2$ dependence, hence lower frequencies are more strongly absorbed, (See Section 1.1.1). As a result, higher frequencies are required to overcome D layer losses. Such an increase in frequency, however, may not always be possible since it produces a change in barrier range.

The resulting absorption occurs generally in a band that is near the equatorward edge of the nighttime auroral oval; with increasing activity the band extends to increasingly later times at nearly constant magnetic latitude. The band is shown schematically as the cross-hatching in Figure 40. Although there is generally weak precipitation continually present in this band, the intense occurrences that often strongly affect HF propagation are highly sporadic, having typical durations of 15 minutes near midnight, an hour or more in the morning and at later times. The shorter events near midnight are related directly to the occurrence of auroral substorms. The longer lasting events in the morning and daytime occur subsequent to substorms as the particles drift in the magnetosphere from midnight through morning to noon. Thus absorption events occur within this band, but are not continuously present under quiet and moderate conditions.

With increasing activity the absorption band expands equatorward (like the auroral oval in Figure 36), absorption events increase in frequency, and the local time extent increases from midnight through sunrise to noon. During magnetospheric storms the band can extend well into the afternoon and evening sectors so as to form a nearly circular ring at all MLT.

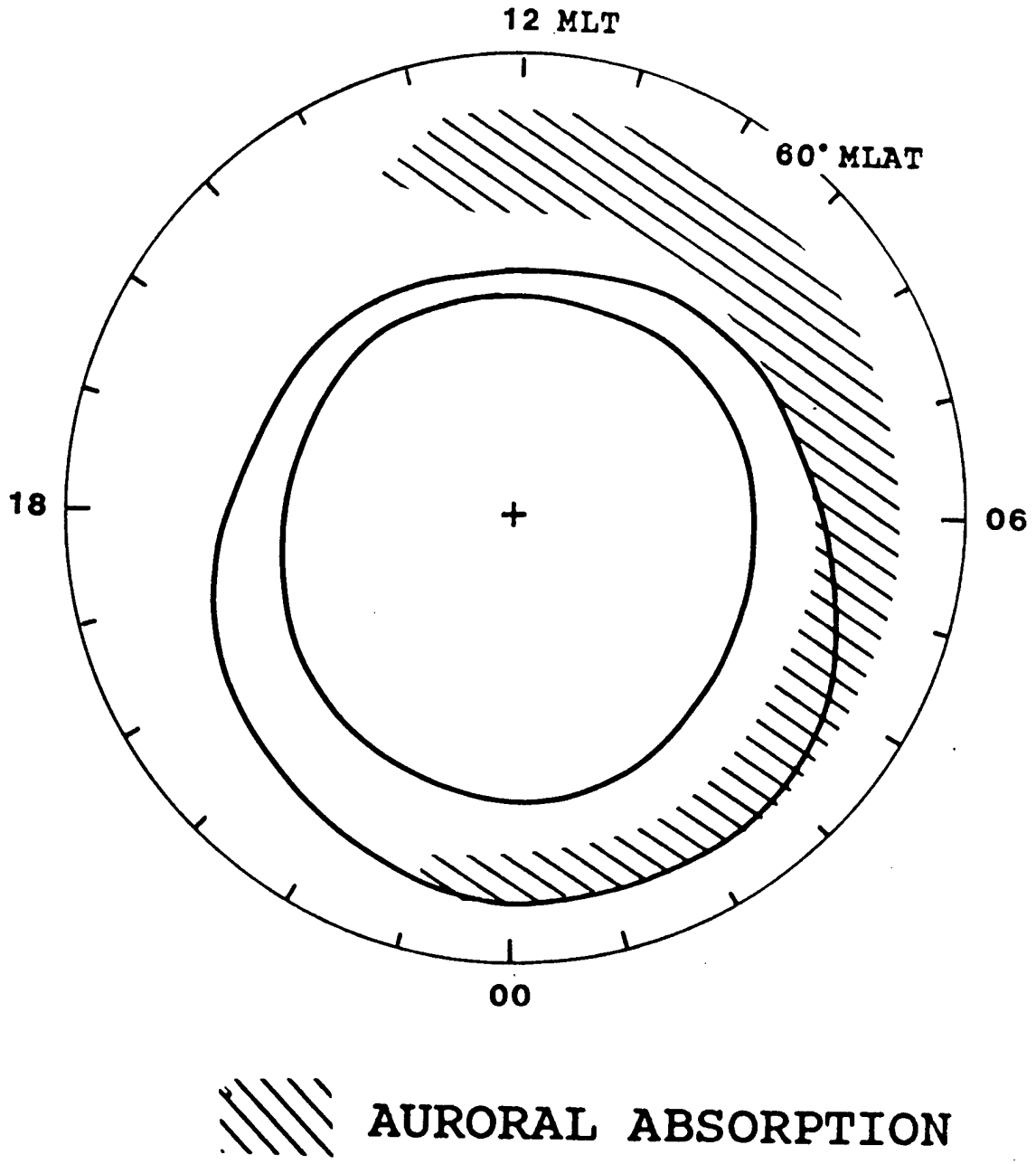


Figure 40. Location of auroral absorption. The region of precipitation of >40 keV electrons is shown schematically in relation to the auroral oval. These electrons ionize principally in the D region and cause absorption preferentially at low frequencies.

A case in which an aircraft is located beneath a region of auroral D ionization so that the down-coming ray is effectively absorbed is illustrated in Figure 41. In such a case the absorption layer results in a high LUF. If, in addition, the midpoint reflection is in the trough where low density plasma results in a low MUF, forcing the radar to low operating frequencies, radar operation may be seriously limited. This case will be discussed further in relation to the trough.

3.2.2 IONOSPHERIC- MAGNETOSPHERIC CONVECTION AND ITS EFFECTS

The second principal magnetospheric phenomenon affecting the high latitude ionosphere is convection of F layer plasma. It arises from an electrical potential generated by the interaction of the solar wind with the earth's magnetic field. The potential has a magnitude of several tens of kilovolts under moderate to strongly active geomagnetic conditions, and is applied with a dawn to dusk polarity across the polar cap, the region enclosed by the auroral oval. The resulting electric field produces a force on the plasma in the direction perpendicular to both the electric(E) and magnetic(B) fields, having the magnitude and direction of the vector **E** x **B**. This force causes plasma to flow across the polar cap from noon to midnight, with return flow to the dayside along the morning and evening flanks of the oval. As a result a circulation pattern is formed consisting of 2 vortex-like cells, a dawn cell in which the circulation is counterclockwise and a dusk cell in which the circulation is clockwise. The domain of convection is shown in relation to the oval as the shaded region with arrows depicting schematically the direction of plasma flow in Figure 42.

Convection produces two major impediments to OTH radar operation: the formation and transport of F region irregularities into the coverage area resulting in Doppler clutter; and an extended region of reduced f_0F_2 called the trough. Clutter can occur within the entire shaded region of Figure 42, whereas the trough occupies the region equatorward of the auroral oval.

3.2.2.1 Doppler Clutter

Plasma flow equatorward of the auroral oval reverses direction near midnight, from westward flow before midnight to eastward flow after midnight. As a result, an OTH radar sees, in the afternoon and evening, westward convecting F region irregularities that result in positive

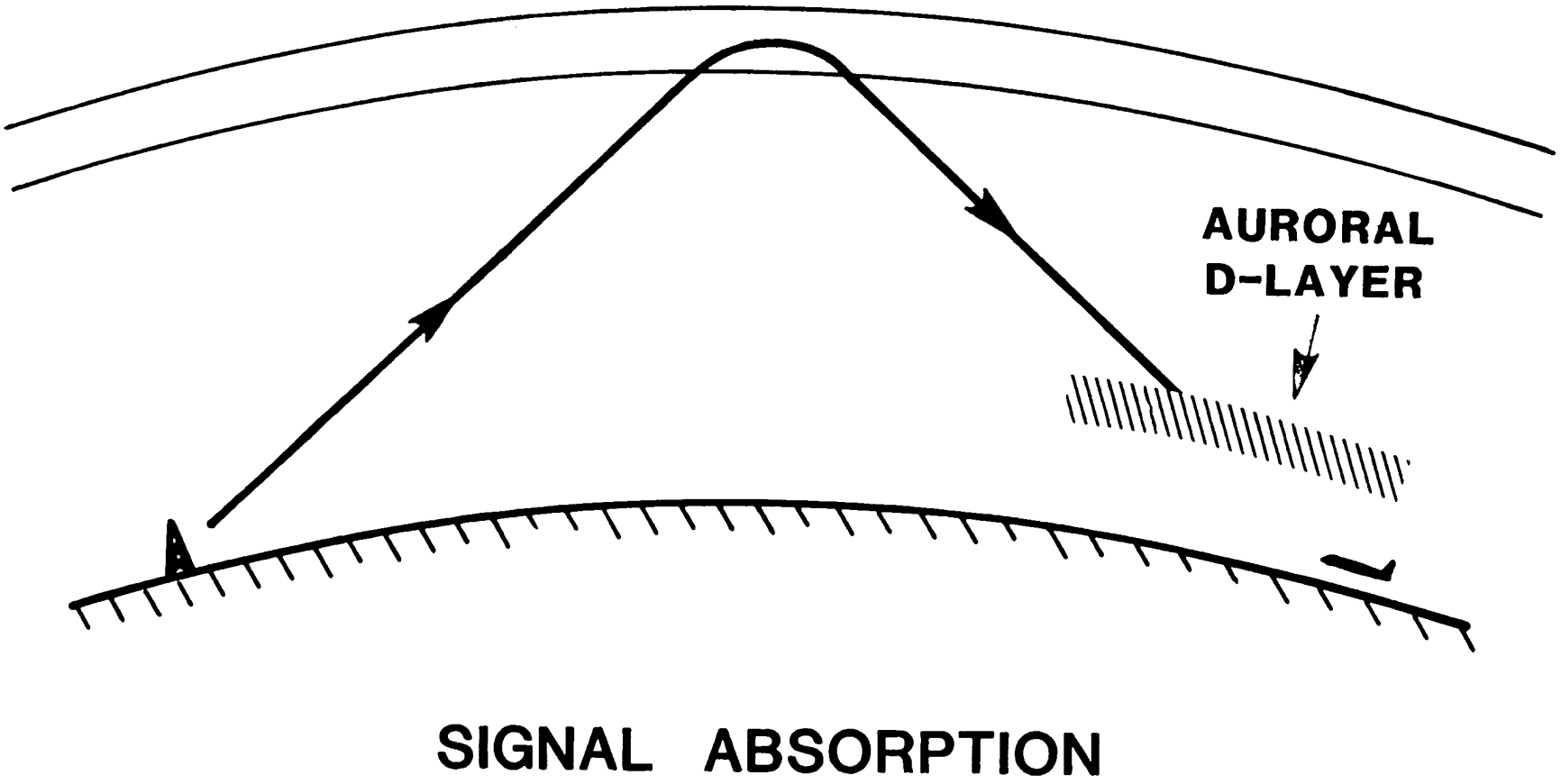


Figure 41. Schematic showing an example of auroral D region ionization absorbing energy from the OTH signal so as to impede detection of the aircraft.

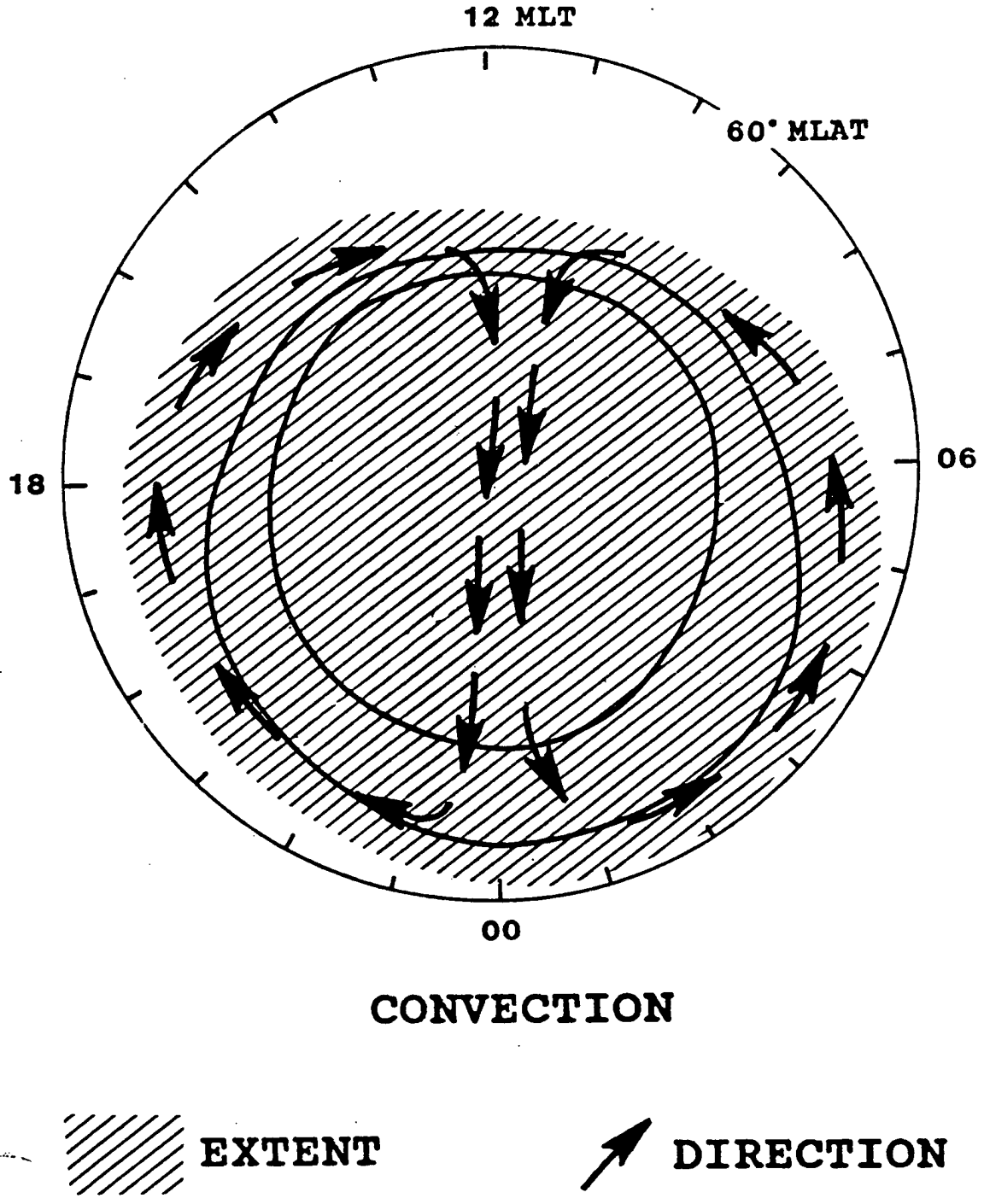


Fig 42. The convection pattern. The arrows schematically indicate the motion of the F-layer plasma within the two vortex-like cells that make up the plasma convection region. The moving F layer irregularities scatter the radar beam to produce Doppler OTH clutter. In addition, a region of low density plasma exists in the trough: the shaded region equatorward of the oval.

Doppler clutter. Near midnight, the flow reverses so that in the morning sector eastward moving irregularities are seen by the radar as negative Doppler clutter. Velocities as high as 1 km /sec are not unusual for the convecting plasma. This process will be illustrated further in Figure 45.

Strong scattering by these irregularities seen line-of-sight by the radar at the same range as the aircraft can produce clutter that masks the aircraft signal. An example of this is shown schematically in Figure 43. The details of scatter geometry and mechanisms, which are somewhat complex, are discussed by G. Sales (in his report PL-TR-92-2123).

3.2.2.2 The F Layer Trough

The trough is formed when low density nighttime plasma is transported by convection towards the terminator and into the daytime, most visibly by the dusk cell. Here high density plasma in the afternoon and evening sectors is replaced to form a region of low density plasma called the F layer trough. The trough occupies the region of convection shown in Figure 42 equatorward of the auroral oval and extends from afternoon through midnight to sunrise. It appears first in the afternoon as a sharp boundary poleward of which f_0F_2 decreases by a factor of 2 to 3 within about 3°, forming a minimum about 2° wide. The trough level decreases with time, falling to nighttime levels near midnight. Low frequencies and spread conditions in the F layer can continue until dawn.

The f_0F_2 within the trough may be so low (2 to 3 MHz during solar minimum, 3 to 4 MHz during solar maximum) that the OTH radar is required to work at the low end of its frequency range. In addition the F layer height is increased within the trough. This condition results in ionospheric tilts at both edges of the trough that can lead to severe off-great-circle propagation.

Another effect of convection occurs where high density plasma is transported from the sunlit daytime ionosphere into and across the polar cap. In the polar cap the convected high density plasma is in the form of patches with dimensions of 100 to 1000 km. These patches develop internal irregularities as they pass across the polar cap. These are transported across auroral arcs to form "boundary features" south of the auroral oval and are sources of intense ionospheric clutter.

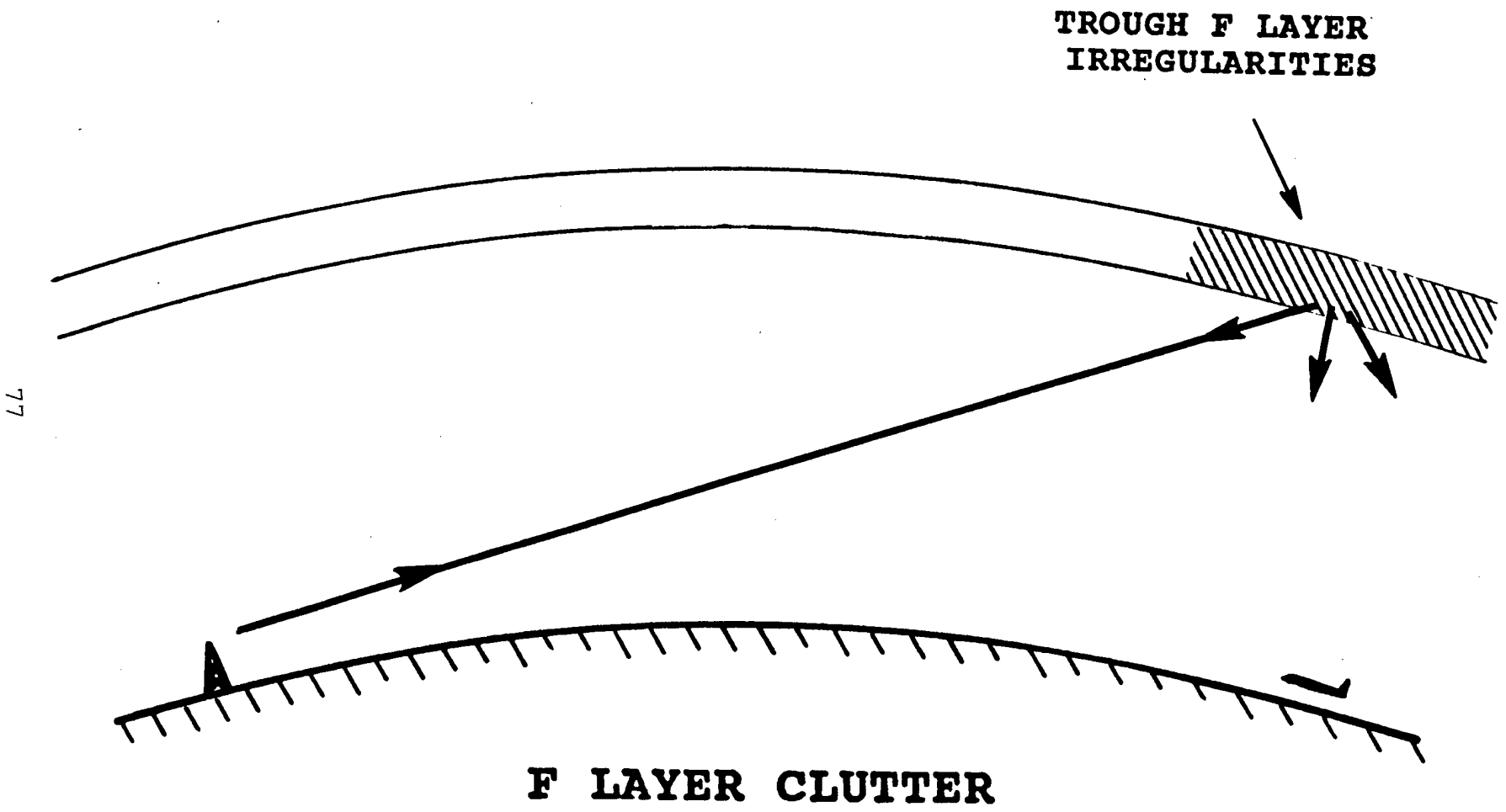


Figure 43. An example of F layer irregularities transported by convection that can scatter to produce clutter at the range of the aircraft.

One impediment to OTH operation occurs in the trough when the radar frequency is too high to be refracted back to earth by the low density plasma of the trough. As a result the signal escapes as shown schematically in Figure 44. The remedy of going to a lower frequency may not be possible if the radar is already operating at its lowest frequency or if there is an enhanced D layer (as shown in Figure 41) requiring the higher frequency to overcome signal losses due to absorption.

3.2.2.3 Composite and Overlay of Auroral and Convective Effects

This section will discuss how the auroral and convective phenomena, described separately in the foregoing, affect in their totality the northernmost ECRS and WCRS OTH segments as a function of time. The regions of auroral oval, auroral E, auroral (D) absorption, and F region convection and trough are plotted together in GMLAT/ GMLT in Figure 45.

The diurnal behavior of this composite is shown in Figure 46 at the same four UT intervals for a Q=4 (disturbed magnetic conditions) as was shown for the oval in Figure 35. Some of the significant effects on the ECRS Segment 1 (northeast looking) during winter are as follows:

18 UT: The trough is in the northeast corner producing positive Doppler clutter.

00 UT: Trough, oval, E layer and polar cap all produce positive Doppler clutter over the northern half of the coverage area. The low f_oF_2 in the trough requires low operating frequencies in the northern two to four sectors of Segment 1.

06 UT: Convection has reversed now producing negative Doppler clutter over much of the coverage area. Absorption occurs sporadically giving rise to contradictory requirements: operating at the lowest frequencies lest the signal escape at its midpoint in the trough (Figure 45); and operating at higher frequencies to overcome signal loss by absorption in the D layer (Figure 41).

The WCRS Segment 3 (northwest looking) is relatively unaffected, some clutter occurring in the northeast corner of the (06 UT) coverage area from 06 UT to 12 UT. Note that the before midnight Doppler clutter is negative, reversing after midnight (12 UT) to positive, the opposite of the east coast behavior.

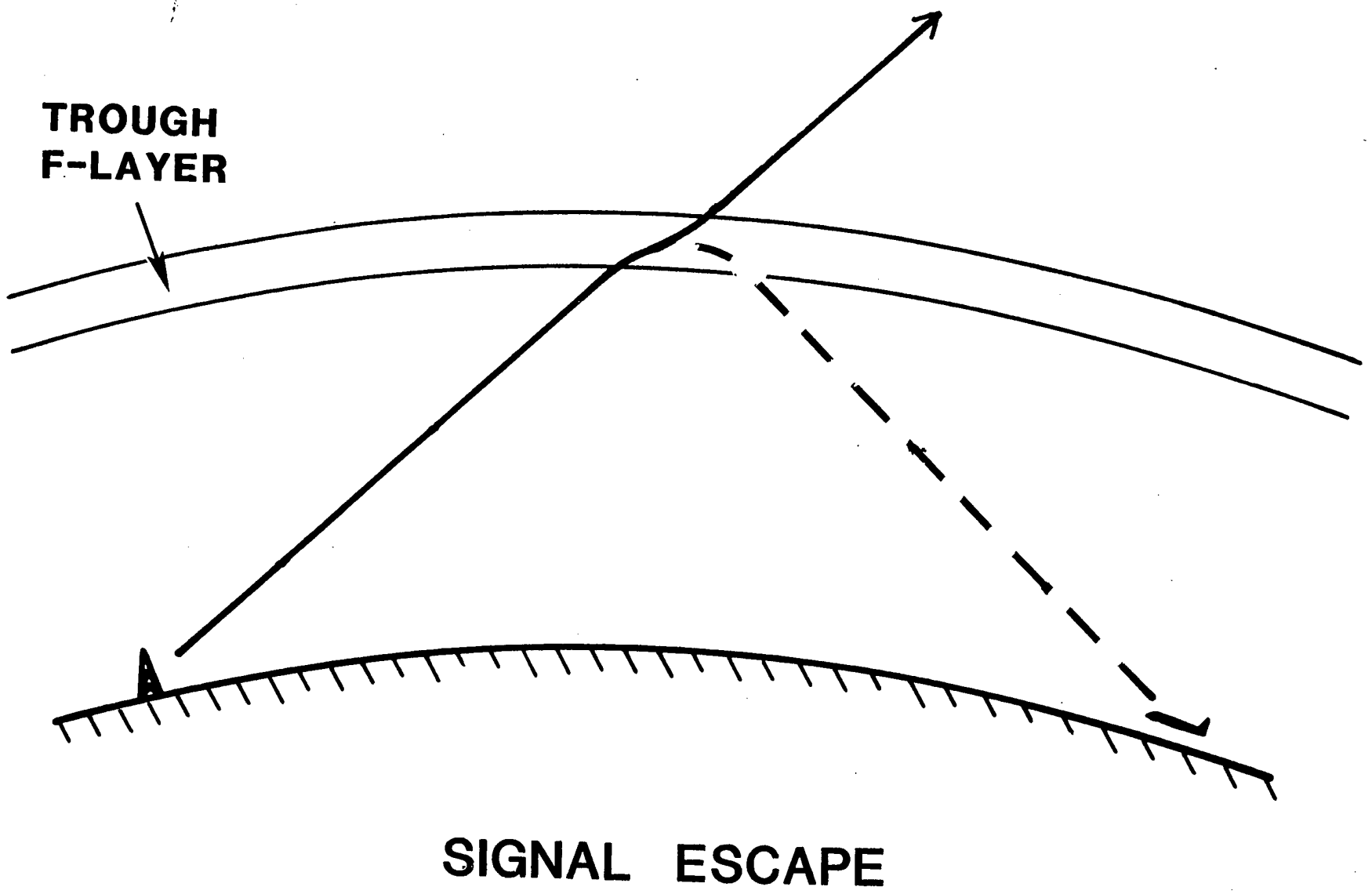


Figure 44. An example of an OTH signal escaping because it has too high a frequency to be refracted back to earth by an F layer of low plasma density. Such an ionosphere is characteristic of the trough.

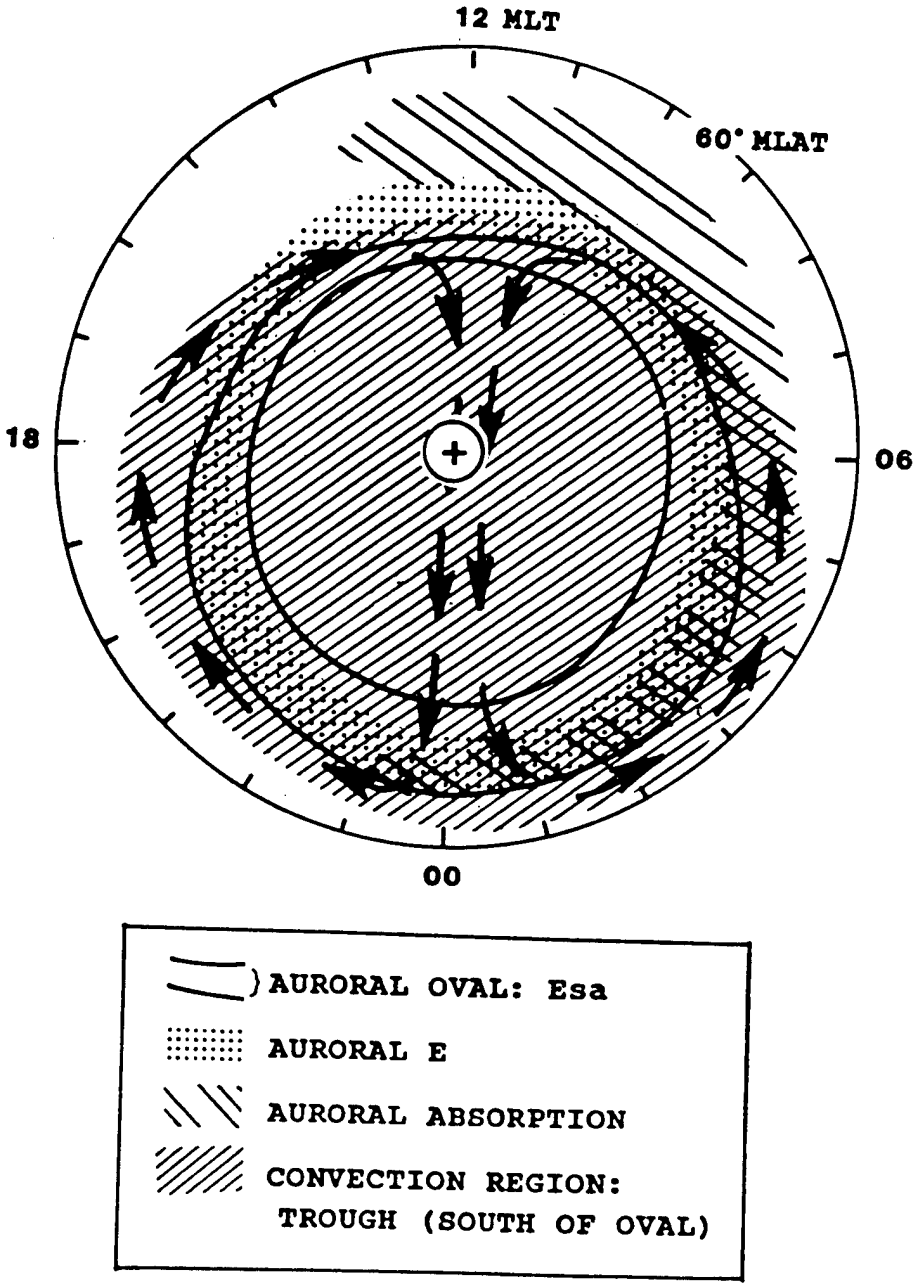


Figure 45. The totality of the high latitude ionospheric phenomena that affect the OTH system as described in this section.

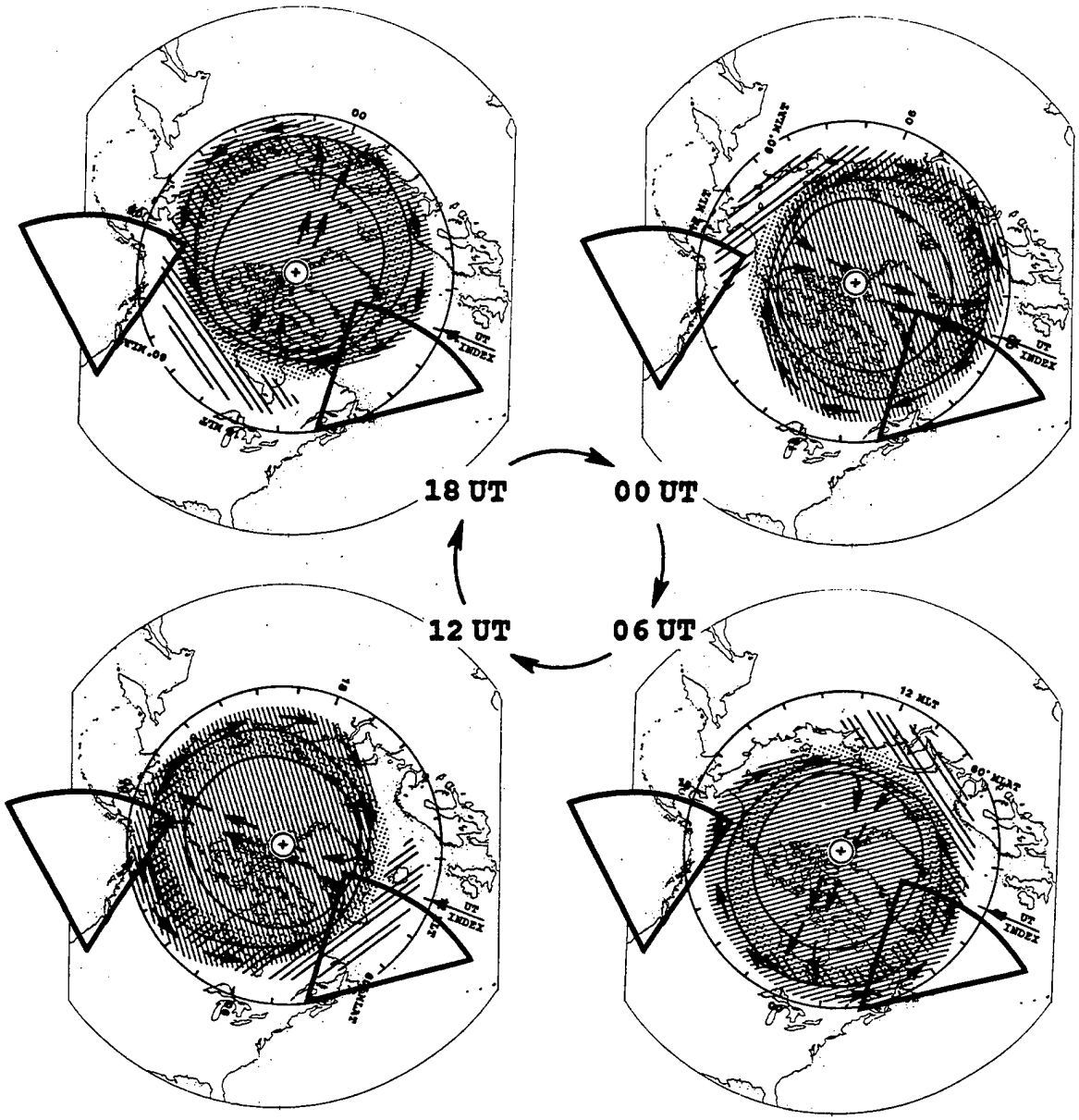


Figure 46. The composite of phenomena shown in Figure 45 superimposed on the northern hemisphere and coverage areas at four UTs.

Figure 46 shows the very extended nature of the various clutter regions. Often intense irregularities extending in azimuth and range provide a strong clutter source: through the side lobes for all segments; and through the back lobes for the east and southeast looking segments and for the west and southwest looking segments. Therefore even if none of the sources of clutter shown in the figure occur in the look direction of the radar, clutter may still affect the radar operation. In addition to the effect of sidelobes, the clutter at long ranges does get folded into the primary ranges of the radar. Thus one sees the equatorial clutter in the northern looking beams of ECRS segment 1 and WCRS segment 3 due to the range folding property of the radar.

It is possible to relate the composite phenomena to the coverage areas as in Figure 46, continuously for all UT. This is accomplished using Figure 45, which is designed as an overlay of Figure 47, a map of the northern hemisphere in GMLAT/ GMLONG. This map can be used in conjunction with Figure 44 to project ionospheric phenomena into the coverage areas as a function of UT as described in the text, (Note that Figures 46 and 47 should be of the same size for an overlay) showing the two northernmost coverage areas. The procedure is as follows:

1. Make a transparent copy of Figure 45, of the composite phenomena.

2. Make an opaque copy of Figure 47, of the map of the showing the two northernmost coverage areas.

3. Place the composite transparency on top of the map so that the magnetic poles coincide. The magnetic pole is the circled cross (+) at the center of each figure.

4. Place a pin or tack through the superimposed crosses so that it forms an axis about which the transparent overlay can rotate over the stationary map.

5. Rotate the overlay to the desired UT as indicated by the arrow on the "UT INDEX". Four UTs are shown in Figure 46, (The UT INDEX is located on the magnetic meridian where UT and MLT are the same. This is analogous to the Greenwich meridian, the geographic meridian on which UT and local time are the same).

The above composite is representative of typical geomagnetically disturbed conditions, although they can vary considerably, often in a very short time. Generally all the phenomena are tied to the auroral oval, which contracts poleward during quiet conditions and expands equatorward during disturbed conditions (Figure 36). As activity increases, the severity of the

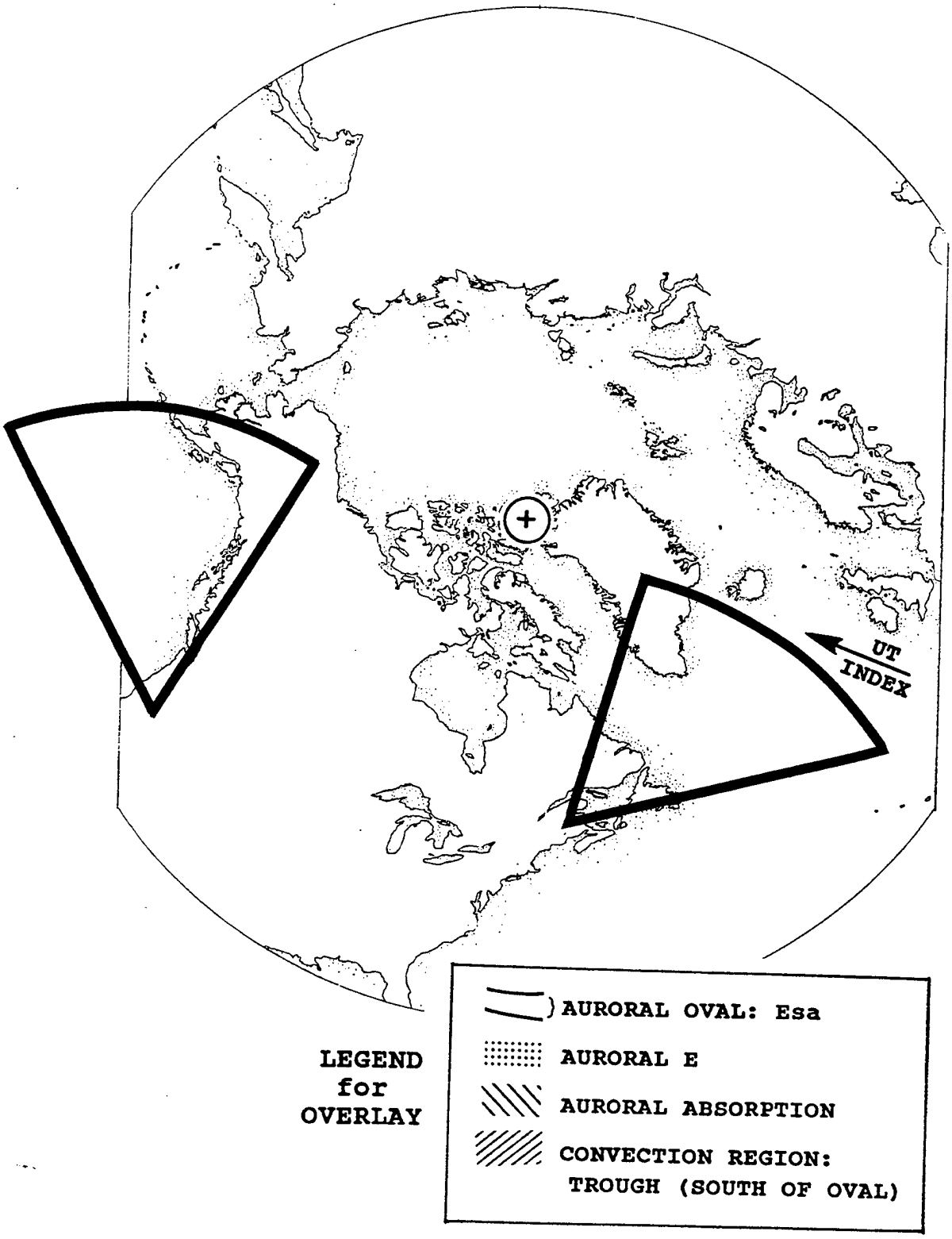


Figure 47. Map of the northern hemisphere including coverage areas in GMLAT/GMLONG. This map can be used in conjunction with Figure 45 to project ionospheric phenomena into the coverage areas as a function of UT as described in the text, (Note that Figures 45 and 47 should be of the same size for an overlay).

effects not only increases, their extent and duration increase: the phenomena enter the coverage area earlier in the afternoon, fill a greater region during the nighttime, and depart at a later time in the morning.

The two types of activity described here, substorms and storms, are discussed in more detail in Section 2.4.

Substorms: One common type of disturbance is known as the substorm, a period of activation and expansion equatorward of the oval as well as the other phenomena. Onset is often within a few minutes of a solar flux ejection, and the disturbance can last for an hour or more. During very quiet periods there may be no substorms although when they do occur, recurrence at intervals of several hours is not unusual.

Geomagnetic Storms: Magnetospheric or geomagnetic storms resulting from solar flares and other disturbances on the sun can expand the composite auroral and convective phenomena into the midlatitudes and there produce chaotic ionospheric conditions especially at night. Section 2 gives a detailed description of geomagnetic storm behavior and its relation to the solar cycle.

3.3 Equatorial Ionosphere

Section 1 contains a passing reference to phenomena like sporadic E and the equatorial anomaly. As the equatorial ionosphere belongs to the category of disturbed ionosphere we will take a detailed look at these features.

3.3.1 SPORADIC E (E_{SQ}) LAYER

In Figure 34 we have seen the diurnal, seasonal and spatial dependence of the E layer. This leads to an intense eastward current (E_{sq} s-solar, q -quiet), that flows by day, over a narrow latitudinal strip along the magnetic equator and is called the equatorial electrojet. This current is the result of the dynamo action of the horizontal wind system and the electrical conductivity of the ionosphere due to the presence of the ions and electrons. The electrojet current is of the order of 10,000 amperes. The electrojet is more intense in local summer. The solar cycle dependence is stronger. Whereas the E layer ionization increases by 50 percent from solar minimum to solar maximum, the E_{sq} strength increases by 100 percent because of the in increase in the wind speed

responsible for the electrojet. The irregularities in the electrojet produce strong sporadic E patches (E_s or E_{sq}). Depending on the geometry of the raypath the strong sporadic E can produce radar clutter as seen in Figure 39 or can support communication to large distances via a two-hop E mode.

3.3.2 F LAYER EQUATORIAL ANOMALY

Because the ionosphere is produced by the ionization of atmospheric gases one would expect the largest concentration of the ionosphere over the equator where the incident solar radiation is strongest on the average during the equinoxes. But due to heating, the ionosphere rises to higher altitudes and flows along the magnetic field lines. Due to gravity it comes down while moving along the field lines. This produces regions of increased concentrations at 15-20° latitude on both the sides of the equator. A schematic of this process is shown in Figure 48. The effect of reduced concentration at equator and enhanced concentration on both sides away from the equator is called the Appleton anomaly. The anomaly has been observed by instruments aboard satellites as shown in Figure 49. The figure shows a single peak at 700 km. At lower altitudes each cut shows two peaks. The peaks are separated more in latitude for cuts at decreasing altitudes, demonstrating the fountain effect. Thus in the given example the ionization is raised to an altitude of 700 km and is separated in latitude by 18° at an altitude cut at 470 km, below which the data are absent. Because the figure is in the geographic coordinate system, the peaks are very asymmetric in latitude. There is a small asymmetry in the peaks with reference to the magnetic equator.

The anomaly is most pronounced at 1400 LT and declines steadily until it disappears around 2000 LT. During high sunspot periods the anomaly redevelops at late evening hours. The latitudinal asymmetry has been explained as due to differences in hemispheric neutral winds and the variation of magnetic declination with latitude.

To the OTH radar ray paths crossing the equatorial region, the equatorial ionosphere has a domelike shape with F layer altitudes higher at the equatorial region. The dome increases the range for these raypaths.

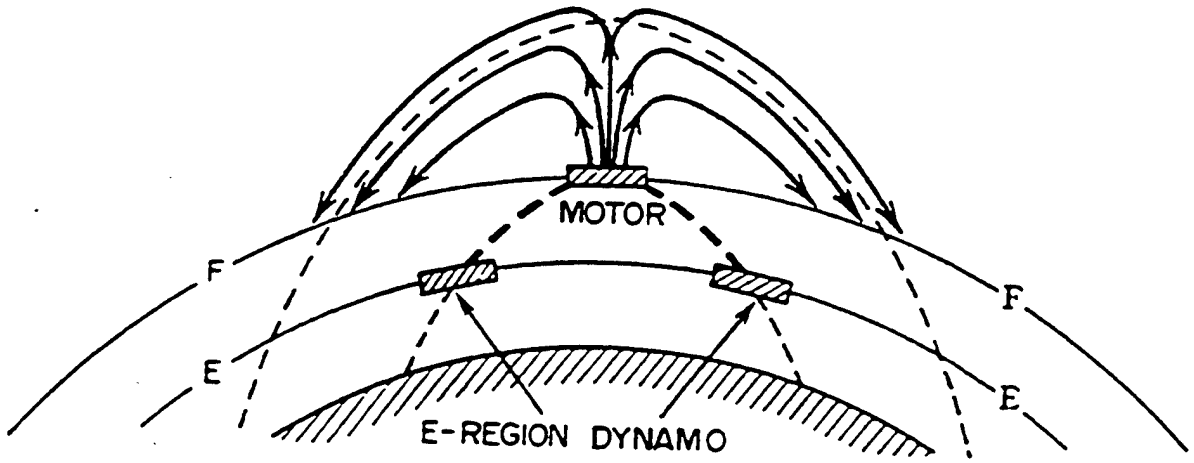


Figure 48. Schematic of F region equatorial anomaly, (Jursa[1] p.9-8). Near the equator the electric fields of the atmospheric dynamo in the E layer are conveyed upwards along the geomagnetic lines of force to the motor in the F layer where they produce an upward movement of the plasma during the day. The raised plasma then diffuses down lines of force to produce enhanced concentration at places on each side of the equator and decreased concentration at the equator itself.

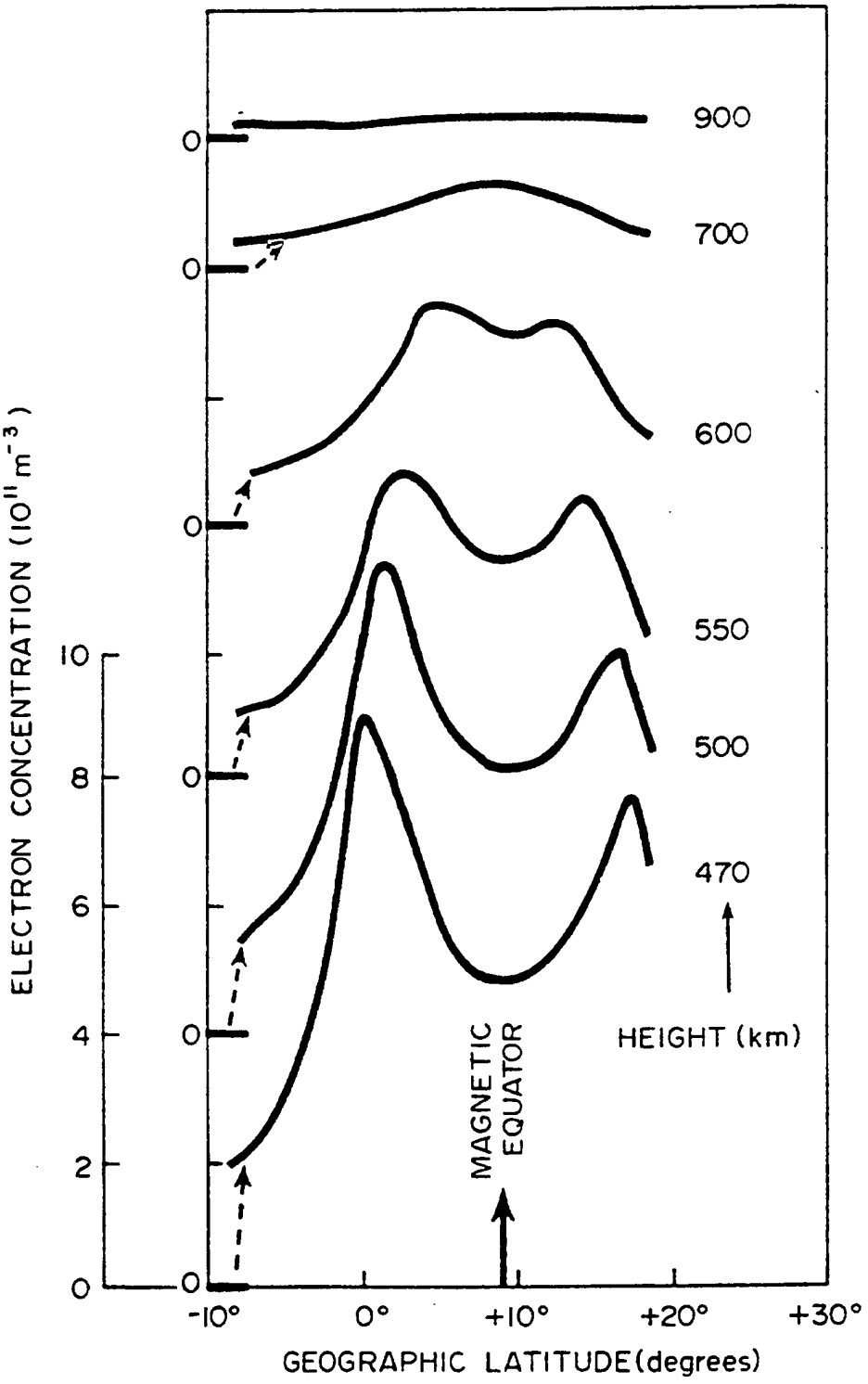


Figure 49. Satellite based observations of F region equatorial anomaly, (Jursa[1] p. 9-9).

The other phenomenon is the equatorial Spread F. It shows spread both in range and frequency. Post sunset equatorial Spread F shows irregularities with a very wide range of scale sizes, 5-6 orders of magnitude, from hundreds of kilometers to tens of centimeters. The rocket and satellite in situ measurements have shown that these irregularities are associated with biteouts in ion concentrations. These biteouts are due to vertical velocities of the ions producing bubbles with holes at the bottom. To the incoherent backscatter radars these bubbles appear like plumes. The plumes are due to vertically rising bubbles and their wakes. These bubbles/ plumes produce strong scintillations on VHF circuits crossing these regions. Figure 50 shows the plumes observed by the Jicamarca incoherent radar and the scintillations on 137 MHz signals received at Ancon from the GOES-1 satellite.

The equatorial ionosphere during periods when the anomaly is present, causes the radar 2 or 3 hop signals to be significantly spread in Doppler. These spread Doppler clutter signals may be of such amplitude that they mask true signals from long ranges. A possible mitigating technique is a use of a different waveform repetition frequency (WRF) to unfold the spread Doppler clutter to different ranges in the ARD data.

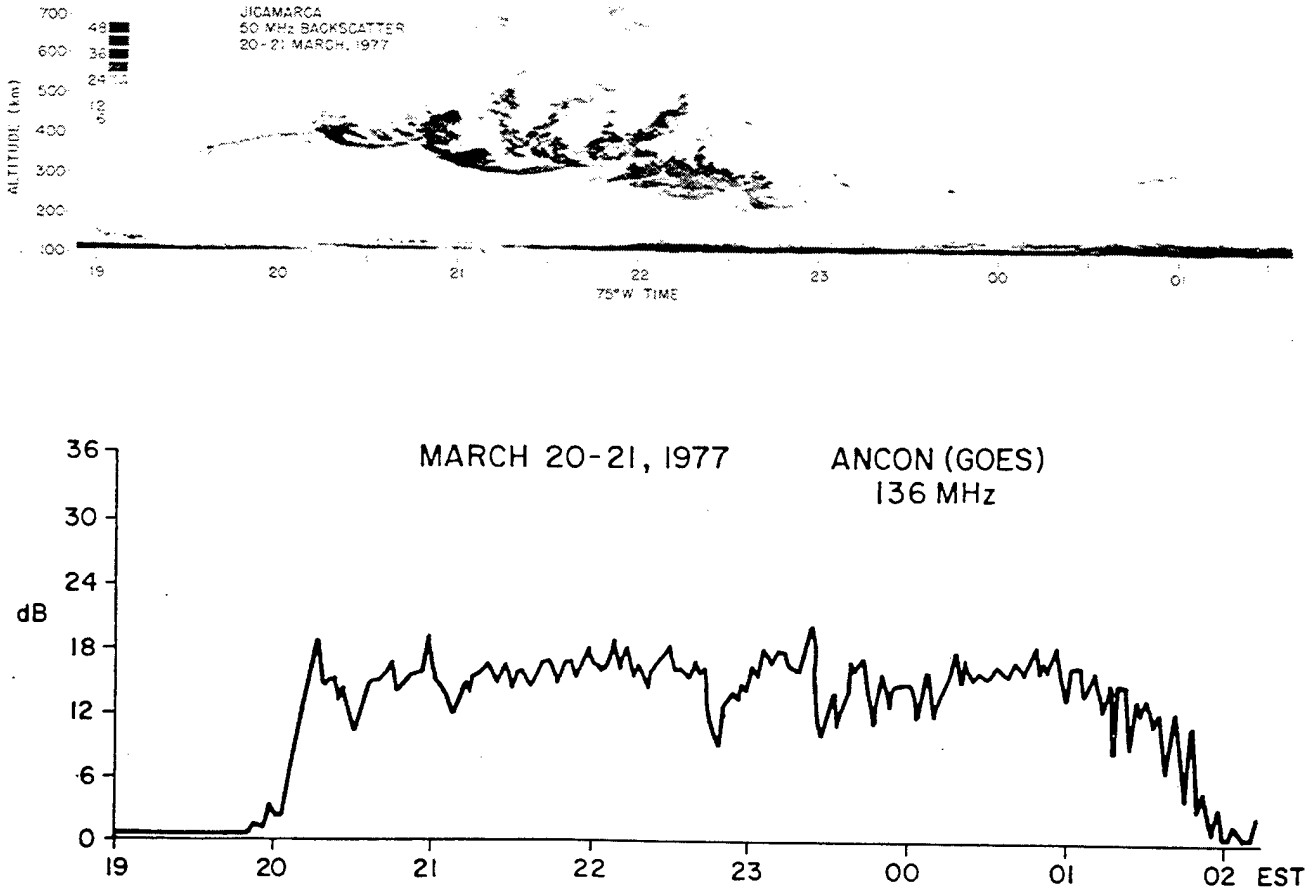


Figure 9-9. Temporal variation of range and intensity (different grey tones) of 50 MHz backscattered power at Jicamarca on 20–21 March 1977 (top panel) and 137 MHz scintillations (bottom panel) over a nearly common ionospheric volume [Basu et al., 1980].

9-10

Figure 50. Observations of equatorial plumes and scintillations produced by the plumes, (Basu et al[9], p.4221). Simultaneous Jicarmarca backscatter power enhancement and VHF scintillation intensity are seen due to irregularities occurring in the same scattering volume.

References

1. Jursa A. S., editor, (1985), *Handbook of Geophysics and the Space Environment,* Air Force Geophysics Laboratory, Air Force Systems Command, USAF, ADA 167000.

2. Roberts W. M. and Rosich R. K, (1971), *Ionospheric Predictions,* vol.3, Telecommunications Research and Engineering Report 13, OT/TRER 13, U. S. Department of Commerce, Office of Telecommunications, Institute for Telecommunication Sciences.

3. Davies K., (1965), *Ionospheric Radio Propagation,* National Bureau of Standards Monograph 80.

4. Pike C. P., (1975), *Defense Meteorological Space Program, Auroral Ionospheric Interpretation Guide*, AFCRL-TR-75-0191, Air Force Surveys in Geophysics, 306, ADA013165.

5. Chapman S., and Bartels J., (1940), *Geomagnetism* vol. II, Oxford University Press, Amen House, London, E.C.4.

6.. Haymes R. C. , (1971), *Introduction to Space Science*, John Wiley & Sons, Inc.,

7. Prochaska, R. D., Capt., (1980), *Geomagnetic Index Calculation and Use at AFGWC*/TN-80/002.

8. Feldstein Y. I. and Starkov G.V., (1967), Dynamics of the auroral belt and polar geomagnetic disturbances, *Planet. Space. Sci.*, **15**, 209.

9. Dandekar B. S., (1979), *Magnetic Disturbance Statistics from a Single Station Q Index Applied to an Actual OTH Radar Situation*, AFGL-TR-79-0296, Environmental Research Papers, No. 687, ADA084808.

10. Basu S., McClure J. P., Basu Su., Hanson W. B., and Aarons J., (1978), Coordinated Study of Equatorial Scintillation and In Situ and Radar Observations of Nighttime F Region Irregularities, *J. Geophys. Res.* **83**, 4219.

Bibliography

Source Book of the Solar-Geophysical Environment, Space Environmental Course, third edition, 1982, AFGWC/WSE, Offut AFB, NE 68113.

Sales, G. S., (1992), *High frequency (HF) Radiowave Propagation*, PL-TR-92-2123, (OTH Handbook Chapter 4) , ADA261726.

Sales G. S., (1992), *OTH Radar System: System Summary*, PL-TR-92-2134, (OTH Handbook Chapter 2), ADA261727.

Dandekar B. S. and Buchau J., Editors, (1995), *Glossary for Over-the- Horizon Backscatter Radars*, PL-TR-95-2127, (OTH Handbook Chapter 6) .